BDEDM 2023

Proceedings of the 2nd International Conference on Big Data Economy and Digital Management

Changsha, China
January 6-8, 2023

EDITORS

Misra Anuranjan
Ke Yan
Wang Yan
Proceedings of the 2nd International Conference on Big Data Economy and Digital Management

January 6-8, 2023, Changsha, China

BDEDM 2023

General Chairs

Professor Yan Wang
Xi’an Peihua University, China

Technical Programme Chair

Professor Ramayah Thurasamy, Universiti Sains Malaysia, Malaysia

Professor Li Yuan, Tongji University, China
Preface

This volume is dedicated to the 2nd International Conference on Big Data Economy and Digital Management (BDEDM 2023) supported by University Malaysia Sabah, Malaysia, held on 6th–8th January 2023 in Changsha, China (virtual conference). The Conference, focused exclusively on recent trends of research in Big Data Economy and Digital Management, attracted about 150 participants around the world. The ultimate plan is to make it a regular event every year based on the very positive response we received which was beyond our expectation.

The immediate purpose of this Conference was to bring together experienced as well as young scientists who are interested in working actively on various aspects of Big Data Economy and Digital Management. The keynote speeches addressed major theoretical issues, current and forthcoming observational data as well as upcoming ideas in both theoretical and observational sectors. Keeping in mind the “academic exchange first” approach, the lectures were arranged in such a way that the young researchers had ample scope to interact with the stalwarts who are internationally leading experts in their respective fields of research.

Besides the invited lectures, a good proportion of the participants also presented their work through contributory talks and posters on this big platform. This was particularly encouraging and of benefit to the young participants, given that there were a good number of scientists of international repute among the participants, the feedback from whom could guide them in the right direction. All the contributions were refereed by experts. This set a standard of its own.

The major topics covered in the Conference are: Big Data in Enterprise Performance Management, Enterprise Management Modernization, Intelligent Management System, Performance Evaluation and Modeling Applications, Enterprise Technology Innovation, etc..

We are indebted to the University Malaysia Sabah for providing us with generous support to organize such a huge conference. We thank all the members of the Organizing Committee who contributed their hard labour to make the Conference a great success. We gratefully acknowledge the experts and reviewers for their valuable suggestions on the review process of the papers submitted.

We sincerely thank the EAI and its staff for the publication of this issue. Last but not the least, we thank all the speakers and participants without whom the Program would not have been such as success. We hope we will have your active participation in the future versions of the Conference as well.

The Committee of BDEDM 2023
Conference Organization

Conference General Chair
Yan Wang, Xi’an Peihua University, China

Publication Chairs
Anuranjan Misra, Greater Noida Institute of Technology, India
Boon Leing TAN, PSB Academy, Singapore
Yan Ke, Wuhan University of Science & Technology, China

Technical Program Committee Chairs
Ramayah Thurasamy, Universiti Sains Malaysia, Malaysia
Li Yuan, Tongji University, China

Local Organizing Chairs
Haocun Wu, South China University of Technology, China
Yan Li, Jianghan University, China

Committee members
Gary I. Lilienthal, Tashkent State University of Law, India
MinHou, Zhejiang Gongshang University, China
Yuriy Shvets, Financial University under the Government of Russian Federation, Russia
GAIKAR VILAS B., University of Mumbai, India
MesfinYemer Yasin, Dire Dawa University, Ethiopia
Amaghousa, Jabrane, Cadi Ayyad University, Morocco
Muhammad Junaid, The University of Haripur, Pakistan
J. Jose Prabhu, Bharathiar University, India

Technical Program Committee
Bashlay Sergiy, Sumy National Agricultural University, Ukraine
Nadeem Malik, University of Balochistan Quetta, France
Ranjit Kr. Sarmah, Assam Down Town University, India
Gabriel Ayodeji Ogunnola, Machine Building Institute Andijan, Uzbekistan
Ferdous Ahmed, International University of Business Agriculture & Technology, USA
Agha Amad Nabi, Dow University of Health and Sciences, Pakistan
Chukwunonso Kelvin Oraedu, University of Nigeria, Nigeria
Muhammad Qasim Shabbir, Mirpur University of Science & Technology, Pakistan
Reza Lotfi, Yazd University, Iran
Nor Yasmin Binti Mhd Bani, Universiti Putra Malaysia, Malaysia
## Contents

**Enterprise Data Management and Economic Statistics**

- Research on Social Responsibility Governance of Platform Enterprises Based on Algorithmic Embedding: A Case Study of Meituan  
  *Yihui Zhang*  
  1
- “Avoid a Crisis” or “Out of the Crisis”: A Study Based on the Impact of Manager Competence on Organizational Resilience of Small and Medium-Sized Enterprises  
  *Mengsu Xu, Zilin Liu, Yiwei Wang*  
  11
- Research on the Influence Mechanism of R&D Investment and Enterprise Performance Based on FEM Algorithm  
  *Juanjuan Liang, Linrong Wu*  
  23
- Research on the Transformation of Financial Accounting to Management Accounting under the Background of Big Data Combined with Algorithm Model  
  *Chao Li*  
  36
- Design and Implementation of Visualization Application Platform of Enterprise Operation Data Based on R Language  
  *He Zhang*  
  45
- Design and Application of Enterprise Intelligent Marketing System Under the Background of Big Data  
  *Benyu Zhao, Qian Chen, Yan Hu*  
  53
- Value Co-Creation of Cultural Tourism Enterprises’ Service Ecosystem in Digital Transformation: -- Case Study if Digital Shaanxi Tourism Group  
  *Lina Sui, Kailun Guan*  
  61
- The Correlation Analysis Between the Informatization Level and the Profit Growth Rate of Small and Medi-um-Sized Enterprises  
  *Xingcheng Guo*  
  74
- Asset Pricing in China’s Stock Market  
  *Lu Yu*  
  83
- The Development and Application of Enterprise Asset Lifecycle Management System Based on Big Data Technology  
  *Jiaying Wang*  
  93
- Evaluation of Employee Cultural Differences in Sino-US Joint Ventures -- Based on IBM SPSS Data Analysis  
  *Jintao Wu, Xinlong Ding*  
  101
- Discuss the Financial Strategy Management of Technology Companies  
  *Meiying Cheng*  
  110
- Design and Construction of Enterprise Financial Data Sharing Service Center Based on Big Data Technology  
  *Yuxin Wang*  
  121
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research on the Influence of Digital Capability on Service Transformation of Manufacturing Enterprises</td>
<td>129</td>
</tr>
<tr>
<td><em>Hailin Dong, Juhong Chen</em></td>
<td></td>
</tr>
<tr>
<td>Research on Synthetic Evaluation Model for Enterprise Integration Development</td>
<td>137</td>
</tr>
<tr>
<td><em>Yuejia Sun, Hongyun Pan, Weixu Liu, Zhaoxi Wang, Lin Lin</em></td>
<td></td>
</tr>
<tr>
<td>Decision Tree-Based Human Resources Forecasting and Enterprise Project Management COVID-19 Impact and Enterprise Response</td>
<td>153</td>
</tr>
<tr>
<td><em>Xue Zhao, Bin Zhao, Zhenjun Du, Fengshan Zou</em></td>
<td></td>
</tr>
<tr>
<td>A Study of Remanufacturing Pricing Decisions Considering Recycling Quality and Retailer CSR Inputs</td>
<td>163</td>
</tr>
<tr>
<td><em>Wensheng Yang, Anxin Hu</em></td>
<td></td>
</tr>
<tr>
<td>Statistical Measurement of Spatial and Temporal Differences in Regional Economic Growth in China</td>
<td>173</td>
</tr>
<tr>
<td><em>Haiyan Wang, Sirong Shang, Zhanni Huang</em></td>
<td></td>
</tr>
<tr>
<td>The Construction of Internal Tax Audit System of Private Enterprises Based on Python</td>
<td>194</td>
</tr>
<tr>
<td><em>He Li</em></td>
<td></td>
</tr>
<tr>
<td>The Moderating Effect of Digitalization on Diversification and Performance of Enterprises</td>
<td>201</td>
</tr>
<tr>
<td><em>Qiong Chen, Zhengmao Yang</em></td>
<td></td>
</tr>
<tr>
<td>Studies on Financial Center Selection Using Spatial Statistical Analysis and Combination Evaluation</td>
<td>211</td>
</tr>
<tr>
<td><em>Longxiao Yuan, Yuxin Jia</em></td>
<td></td>
</tr>
<tr>
<td>Business Strategy and Investment Efficiency</td>
<td>220</td>
</tr>
<tr>
<td><em>Xuetong Ding</em></td>
<td></td>
</tr>
<tr>
<td>An Empirical Study of ESG Performance and Corporate Investment Efficiency – Moderating Effect of External Pressure</td>
<td>232</td>
</tr>
<tr>
<td><em>Chunxia Fu, Zhiguang Guo, Luorao Yang</em></td>
<td></td>
</tr>
<tr>
<td>International Oil Price Uncertainty, Operating Net Cash Flow and Corporate Investment Analysis Based on Stata and Python</td>
<td>242</td>
</tr>
<tr>
<td><em>Yi Wang, Ye Miao, Xu Liang</em></td>
<td></td>
</tr>
<tr>
<td>Design and Implementation of Entrepreneur Evaluation and Analysis System Based on AHP Method</td>
<td>253</td>
</tr>
<tr>
<td><em>Na Guo, Xiaoxian Gong</em></td>
<td></td>
</tr>
<tr>
<td>Constructive Ideas of Comprehensive Transportation Dispatch and Emergency Command System in the New Era</td>
<td>260</td>
</tr>
<tr>
<td><em>Bin Wei, Jiandong Cao, Weirong Luo</em></td>
<td></td>
</tr>
<tr>
<td>Design and Implementation of Lightweight Web Asset Identification System</td>
<td>267</td>
</tr>
<tr>
<td><em>Zhengde Li</em></td>
<td></td>
</tr>
<tr>
<td>Research on Dilemmas and Rescue Measures of Enterprises in the Yangtze River Delta Under the COVID-19 Pandemic</td>
<td>276</td>
</tr>
<tr>
<td><em>Lu Li, Wenyan Zhang, Qing Liu</em></td>
<td></td>
</tr>
<tr>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Discussion on the Development Strategy of Digital Marketing of Agricultural Products from the Perspective of Long Tail Theory Take Agricultural Products with Geographical Indications as an Example</td>
<td>285</td>
</tr>
<tr>
<td>Huiru Wang, Xiaoyuan Yu</td>
<td></td>
</tr>
<tr>
<td>Litian Chen, Xinran Chang, Junhong Zhong, Yufeng Wang</td>
<td></td>
</tr>
<tr>
<td>Research on the Relationship Between Absorptive Capacity and Innovation Capacity of AI enterprises: the Moderating Effect of R&amp;D Leap and Value Cognition Complexity</td>
<td>304</td>
</tr>
<tr>
<td>Litian Chen, Rui Wu, Yufeng Wang, Junhong Zhong</td>
<td></td>
</tr>
<tr>
<td>The Management Rights and Accounting Conservatism – Data Mining and Data Analysis of 4404 Listed Companies</td>
<td>317</td>
</tr>
<tr>
<td>Chunyu Meng</td>
<td></td>
</tr>
<tr>
<td>Statistical Studies on the Influencing Factors of College Students' Mystery Box Consumption Based on Path Analysis: A Case Study of Guangzhou</td>
<td>324</td>
</tr>
<tr>
<td>Hua He, Wangling Feng</td>
<td></td>
</tr>
<tr>
<td>Research on the Realization Model of Accounting Information Sharing</td>
<td>334</td>
</tr>
<tr>
<td>Jilin Qu, Yicai Zhu</td>
<td></td>
</tr>
<tr>
<td>The Role of Information and Communication Technologies in Modern Business Management</td>
<td>347</td>
</tr>
<tr>
<td>Zhian Ye, Natalia Karmina</td>
<td></td>
</tr>
<tr>
<td>Statistical Analysis of the Adaptability of Street-Stall Economic Policies in the Post-Epidemic Era Based on SPSS 22.0-Taking -Taking Zhejiang Province and Shaanxi Province as Examples</td>
<td>356</td>
</tr>
<tr>
<td>Xiayi Pan, Minyu Huang, Shiyu Huang</td>
<td></td>
</tr>
<tr>
<td>Research on Enterprise Value Evaluation of Electronic Manufacturing Industry</td>
<td>376</td>
</tr>
<tr>
<td>Wensheng Yang, Ruoqi Xin</td>
<td></td>
</tr>
<tr>
<td>Survey on the Awareness of Basic Drugs among Urban Residents in Jilin Province –Based on Binary Logistic Regression Analysis</td>
<td>386</td>
</tr>
<tr>
<td>Kaiyue Shen, Yanyin Cui, Jie Chen, Hui Yue, Zheng Xie, Chen Ding, Jingshuo Liu, Biao Zhang</td>
<td></td>
</tr>
<tr>
<td>Chinese Patent Analysis for Foreign Object Detection in Wireless Charging</td>
<td>399</td>
</tr>
<tr>
<td>Qianxiu Liu, Hongshen Pang, Jingdong Tian, Xinghua Wei, Danhui Song</td>
<td></td>
</tr>
<tr>
<td><strong>Big Data Evaluation and Technology Innovation</strong></td>
<td></td>
</tr>
<tr>
<td>Risk Identification of Bulk Commodity Electronic Trading Based on HHM</td>
<td>409</td>
</tr>
<tr>
<td>Xiangping Deng, Xiaoyan Gu</td>
<td></td>
</tr>
<tr>
<td>The Influence of Asymmetry of Positive Feedback Trade on Expected return: An Empirical Study of China Securities 800</td>
<td>419</td>
</tr>
<tr>
<td>Qinyi Yu</td>
<td></td>
</tr>
<tr>
<td>Digital Transformation, Dynamic Capability and Green Technology Innovation: Empirical Evidence Based on Text Analysis Methods</td>
<td>430</td>
</tr>
<tr>
<td>Jiarun He, Binyi Sun</td>
<td></td>
</tr>
<tr>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Big Data Evaluation on Blue Carbon Economic Development Level in Daguang Bay, Jiangmen City</td>
<td>441</td>
</tr>
<tr>
<td>Jianhua Xiao, Wanqing Zhu</td>
<td></td>
</tr>
<tr>
<td>Research on Evaluation of Innovation and Entrepreneurship Ability of Hebei Province Based on GPCA-HCA Model in Big Data Era</td>
<td>450</td>
</tr>
<tr>
<td>Shanshan Wang, Lihong Han</td>
<td></td>
</tr>
<tr>
<td>Chao Li</td>
<td></td>
</tr>
<tr>
<td>Big Data Evaluation on Marine Ecological Civilization in Guangdong Province</td>
<td>471</td>
</tr>
<tr>
<td>Jianhua Xiao, Qiqin Hu</td>
<td></td>
</tr>
<tr>
<td>The Innovation and Application of Big Data Technology in Cross-Border E-Commerce Comprehensive Test Zone in China</td>
<td>480</td>
</tr>
<tr>
<td>Lixia Hu</td>
<td></td>
</tr>
<tr>
<td>Application of Job Analysis in Recruitment Based on Big Data Technology: Take F Company as an Example</td>
<td>489</td>
</tr>
<tr>
<td>Li Hang, Huam Hon Tat</td>
<td></td>
</tr>
<tr>
<td>Jing YIN, Guanghui ZHU, Mingle ZHOU</td>
<td></td>
</tr>
<tr>
<td>Big Data Analysis of the Business Model Innovation</td>
<td>511</td>
</tr>
<tr>
<td>Yanjing Gu</td>
<td></td>
</tr>
<tr>
<td>How Shareholders' Right of Speech Affects the Performance of Green Innovation--Big Data Practice Based on R Language Brucer Package</td>
<td>518</td>
</tr>
<tr>
<td>Yulu Chen, Junzheng Wu, ChuHan Lin</td>
<td></td>
</tr>
<tr>
<td>Dynamic Evaluation of Eco-Economic Development in Prefecture-Level Cities in Yunnan Province</td>
<td>526</td>
</tr>
<tr>
<td>Zhanni Huang, Yunhe Zhao, Hansheng Hu, Haoyang Shen, Sirong Shang</td>
<td></td>
</tr>
<tr>
<td>Construction of Leader Training System Based on Internet Technology and Big Data Thinking</td>
<td>540</td>
</tr>
<tr>
<td>Jin Li, Qingyu Zhao</td>
<td></td>
</tr>
<tr>
<td>A Study on the Application of Digital Technology in Large-scale Events Risk Management</td>
<td>550</td>
</tr>
<tr>
<td>Jiangtian Wang, Jie Shen, Xinrui Chen, Jiankang Zhang</td>
<td></td>
</tr>
<tr>
<td>The Development of Digital Technology Efficiency in the Communication of Large-Scale Events</td>
<td>561</td>
</tr>
<tr>
<td>Jie Shen, Jiangtian Wang, Jiankang Zhang</td>
<td></td>
</tr>
<tr>
<td>Research on the Association of Digital Media Art and Cultural and Creative Industry</td>
<td>573</td>
</tr>
<tr>
<td>Shuyao He, JingHong Chen</td>
<td></td>
</tr>
<tr>
<td>Research on the Dependency Distance of Quantifiers in Modern Chinese Based on Big Data Text Taking the Dot Quantifier “(particle)” as an Example</td>
<td>580</td>
</tr>
<tr>
<td>Linlin Zou</td>
<td></td>
</tr>
</tbody>
</table>
Research on the Impact of Different Power Structures on the Supply Chain of Fresh Produce

Dandan Wang, Wei Gao

A Research on the Development of a Publicly Owned Natural Assets Management and Supervision Platform

Jiaqiang Ren, Huijing Chu

The Spatial Spillover of Digital Economy on Green Innovation: Evidence from China

Yuming Zhai, Ning Xie

Research on E-Commerce Trading System Market-Oriented Transformation Which Based the Value Co-Creation Theory

Ming Shi, Ming Zhang, Yu Duan

Environmental Regulation, Technological Innovation and Industrial Structure Optimization

Xiaoyan Ren, Xiaoqi Qin, Yaning Li

Research on Comprehensive Evaluation of Railway Technical Regulation System Based on FAHP

Qian Li, Junting Lin, Zhuo Cai, Yang Wang

Research on New Media Precision Marketing Method Based on Big Data Information Automatic Push

Yi Zhou, Yu Chen

A Design Method of Logistics Support Simulation Support Software

Yong Sun, Shaopan Zhang, Guangzhao Lu, Jingqing Zhao, Jianhui Tian, Jirong Xue

Design of User Portrait Analysis System Based on E-Commerce Big Data

Yu Chen, Yi Zhou

PE Factor in the Asset Pricing

Jiaxin Lyu

Structural Integration and Optimization of Human Resource Management Under the Background of Artificial Intelligence

Linqiang Zhang, Chang Wang

Novel Blockchain-Based Privacy Protection for Smart Home

Lifeng Pan, Shuguang Zhou

The Influence of Institutional Investor Shareholding on the Innovation Level of Listed Companies

Qingyuxi Yang, Hangqin Xiang

Research on the Competitiveness of China's High-Tech Industry Based on Entropy-TOPSIS Method

Dan Zhao, Hanxin Zhang, Qinya Dai, Yu Bai, Lihuang Chen

A Critical Analysis on the Cost Planning in Building Project Success: A Theoretical Review

Zhenquan Zhou, Deprizon Syamsunur, Xinyi Wang
Features of Digital Payment Systems in China and the Prospects for Their Implementation Around the World
Guoliang Sun, Vladimir Grigoriev

The Use of Information Technology in the Strategy of Internationalization Of Banking
Yi Yao, Vladimir Grigoriev

An Electronic Channel Operation Method for Power Marketing Based on User Behavior Analysis
Qirui Chen, Wei Zhang, Jianing Xu, Jiali Yu, Ying Jiang

Research on Scientific Management of Living Materials During the Coronavirus Disease 2019 Epidemic
Feng Chen

Research on Service Quality of College Express Market from the Perspective of Customers -Take Universities in Nanchang as an Example
Rui Zhang

Promoting Class Safety Benchmarking Evaluation Based on Pearson Correlation Analysis
Xiuhua Li, Lei Fu, Chao Zhang, Shangmin Li, Yao Zhang, Jianing Si

Logistics Management and Data Simulation and Prediction

Literature Review of The Impact of China-ASEAN Free Trade Area on Industrial Structure Based on Citespace -- Concurrently Discussing the Impact of RCEP on China's Industrial Structure Upgrading
Shuimeng Wan, Xinzao Huang

The Influence of the Change of Labor Relations on Willingness to Share in Sharing Economy Based on the Structural Equation Model
Junfeng Liao, Yongyu Liao, Zhengyan Yang, Siyuan Li

Analysis of Exchange Rate Fluctuations between RMB and US Dollar
Zilin Liu, Wentian Wang, Chunlei Zhao

Systematic Risk Prediction in Commercial Banks Based on Random Forest and BP Neural Network
Junbin Zhang, Peiying Zhang, Shiyang Song, Junyu Su, Jinhai Tang

Development Potential Analysis and Countermeasure Research of Prepared Dishes Industry Based on SWOT
Xiaoning Jiang

Analysis of the Relationship Between ETF Volatility and Liquidity Based on ARMA-GARCH Model
Qingyuan Feng

The Gratification Driving User Attitude and Continuance Use of Mobile Payment Services in China
Liming Zhang, Ruochen Hu, Kuankuan Luo

Temporal Fusion Transformers Model for Traffic Flow Prediction
Yuxuan Zhou
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design and Development of University Smart Campus Platform Based on Big Data</td>
<td>851</td>
</tr>
<tr>
<td>Daigen Huang</td>
<td></td>
</tr>
<tr>
<td>Research on Precision Marketing Strategy Based on Clustering Algorithm</td>
<td>860</td>
</tr>
<tr>
<td>Zhengmin Fu</td>
<td></td>
</tr>
<tr>
<td>Improvement of Traditional Bond Default Identification Model Based on ESG Score</td>
<td>867</td>
</tr>
<tr>
<td>Zhiyuan Bai, Nanyang Huang</td>
<td></td>
</tr>
<tr>
<td>Research on Innovative Practice of Financial Management of Power Grid Enterprises Based on RPA Technology</td>
<td>882</td>
</tr>
<tr>
<td>Changping Chen, Yansheng Li, Wenjie Yang, Bing Lin</td>
<td></td>
</tr>
<tr>
<td>Economic Development Prediction Model Based on Deep Convolutional Neural Network</td>
<td>893</td>
</tr>
<tr>
<td>Taisong Liu, Xiaona Cai, Songyu Xie, Shanting Tan, Zihan Ying</td>
<td></td>
</tr>
<tr>
<td>Research on Omni-Channel Supply Chain Structure Based on BOPS Mode</td>
<td>901</td>
</tr>
<tr>
<td>Shuang Zhou</td>
<td></td>
</tr>
<tr>
<td>Corporate Performance Prediction Based on BP Neural Network</td>
<td>909</td>
</tr>
<tr>
<td>Hongyan Ye</td>
<td></td>
</tr>
<tr>
<td>A Hybrid Model Integrating LSTM with Multiple GARCH-Type Models for Volatility and Var Forecast</td>
<td>918</td>
</tr>
<tr>
<td>Jiayi Liu</td>
<td></td>
</tr>
<tr>
<td>Research on BIM Based EPC Project Lifecycle Information Management</td>
<td>930</td>
</tr>
<tr>
<td>Yuzhuo Zhang, Yuxuan Zhu, Feng Guo, Xuhang Han</td>
<td></td>
</tr>
<tr>
<td>Application Design of Road, Bridge and Tunnel Engineering Information Management Based on Blockchain</td>
<td>938</td>
</tr>
<tr>
<td>Ping Zhang, Mingchao Liao</td>
<td></td>
</tr>
<tr>
<td>Research on the Mediation Model and Empirical Analysis for the Influence of Digital Investment on Retail Business Performance</td>
<td>945</td>
</tr>
<tr>
<td>Xing XING, Qinglie WU</td>
<td></td>
</tr>
<tr>
<td>Big Data Used for Accurately Subsidize Research and Practice in University</td>
<td>958</td>
</tr>
<tr>
<td>Bin Xu</td>
<td></td>
</tr>
<tr>
<td>Research on Text Recognition System of Logistics Enterprise Policy Based on Text Mining</td>
<td>966</td>
</tr>
<tr>
<td>Jiahui Wang, Yidi Wang, Limei Xu</td>
<td></td>
</tr>
<tr>
<td>Typical Practice of Digital Management in Electrical Industry: Characteristics and Cases of International Energy Internet Practice</td>
<td>974</td>
</tr>
<tr>
<td>Jing Wang, Yuwei Wang, Jiachen Wang, Guanghui Wu</td>
<td></td>
</tr>
<tr>
<td>China-ASEAN International Logistics Development Strategy</td>
<td>981</td>
</tr>
<tr>
<td>Tian Fang</td>
<td></td>
</tr>
<tr>
<td>Coordination of Intermodal Transport in Yangtze River Economic Belt in China Based on the Fourth-Party Logistics</td>
<td>991</td>
</tr>
<tr>
<td>Xiaofang Chen, Junli Zheng, Junbo Zhao</td>
<td></td>
</tr>
<tr>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Research and Application Analysis of Two-Level Operational Risk Prevention and Control System of “Online State Grid”</td>
<td>1010</td>
</tr>
<tr>
<td><em>Yuling Chen, Wei Zhang, Ying Jiang, Fei Lou, Qirui Chen</em></td>
<td></td>
</tr>
<tr>
<td>Research on Rich-Club Phenomenon in Beijing Urban Area Based on Trajectory Data</td>
<td>1016</td>
</tr>
<tr>
<td><em>Xiaotong Guo</em></td>
<td></td>
</tr>
<tr>
<td>Research on Export Tax Refund Declaration Robot Based on RPA</td>
<td>1022</td>
</tr>
<tr>
<td><em>Yin Ping</em></td>
<td></td>
</tr>
<tr>
<td>Research on the Influence of International Cooperation of ESI Highly Cited Papers Based on CNCI</td>
<td>1033</td>
</tr>
<tr>
<td><em>Ying Tan, Yan Zhang, Mei Tang, Wanxin Zhao</em></td>
<td></td>
</tr>
<tr>
<td>Evaluation Method of Post Competency Based on Fuzzy Analytic Hierarchy Process</td>
<td>1042</td>
</tr>
<tr>
<td><em>Di Cui</em></td>
<td></td>
</tr>
<tr>
<td>Research on Public Participation Model Based on Block-Chain Mode: Taking Sponge City Construction as an Example</td>
<td>1053</td>
</tr>
<tr>
<td><em>Jin Shang</em></td>
<td></td>
</tr>
<tr>
<td>Analysis of the Problems and Policy Suggestions of China’s Macroeconomics Based on the Theory of Short-Term Economic Fluctuation and Numerical Simulation with Matlab</td>
<td>1059</td>
</tr>
<tr>
<td><em>Kexin He, Yunyan Huang, Yujin Zhang</em></td>
<td></td>
</tr>
<tr>
<td>Research on the Dissemination and Evolution of Online Public Opinion on Unconventional Emergencies Based on Social Networks</td>
<td>1069</td>
</tr>
<tr>
<td><em>Mingyu Xu</em></td>
<td></td>
</tr>
<tr>
<td>Cause Analysis of Ship Accidents Based on Chi-Square Test</td>
<td>1078</td>
</tr>
<tr>
<td><em>Xintong Liu, Qing Yu, Yutian Hu, Wenting Lu, Aipan Zhang, Xujiang Bo</em></td>
<td></td>
</tr>
<tr>
<td>Analysis of Southeast Asian Potential Ports in the Container Shipping Network Between China and Southeast Asian Nations</td>
<td>1083</td>
</tr>
<tr>
<td><em>Siheng Chang, Jinhai Chen, Peng Peng</em></td>
<td></td>
</tr>
<tr>
<td>A Visual Analysis of Chinese Internet Healthcare Research Based on Bibliometrics</td>
<td>1093</td>
</tr>
<tr>
<td><em>Yufang He, Kaiyue Shen, Chengye Zhang, Hongjuan Wen</em></td>
<td></td>
</tr>
<tr>
<td>Analysis of College Students’ Consumption Structure and Irrational Consumption Behavior Based on ELES Model</td>
<td>1103</td>
</tr>
<tr>
<td><em>Xucong Hu, Weihong Wang, Sirui Huo, Fengzhi Chen</em></td>
<td></td>
</tr>
<tr>
<td>Analysis of influencing factors of energy data market construction and development based on entropy weight DEMATEL method</td>
<td>1115</td>
</tr>
<tr>
<td><em>ChunLei Zhou, JunNi Li, TianGuang Yang, JinWei Song, Yu Zhang, SiYuan Gao, YaFeng Wen</em></td>
<td></td>
</tr>
</tbody>
</table>
Research on Social Responsibility Governance of Platform Enterprises Based on Algorithmic Embedding:
A Case Study of Meituan

Yihui Zhang
zangyh1013@163.com
Beijing Jiaotong University, Beijing, China

Abstract: The role of platform enterprises in economic and social development is becoming more and more important, but the phenomenon of social responsibility deficiency and alienation is frequently seen. In this paper, on the basis of the three levels of "social responsibility as an independent operating entity", "social responsibility as a commercial operation platform" and "social responsibility as a social resource allocation platform" found in the existing studies, we embed "social responsibility as a platform for algorithm development and operation". Based on the three levels of "social responsibility as a platform for algorithm development and operation", this paper embeds the fourth dimension of "social responsibility as a platform for algorithm development and operation", and makes an innovative definition of social responsibility of platform enterprises, and takes Meituan as an example to conduct a case study. Finally, this paper summarizes algorithm governance into two dimensions, including the rationalization of platform enterprises' disclosure of the potential impact of algorithm sources and decisions and the empowerment of stakeholders' responsibility.

Keywords: Algorithm, Platform Enterprise, Social Responsibility, Meituan.

1 INTRODUCTION

The benefits of digital transformation are becoming increasingly evident in the new wave of technological revolution. Technologies such as big data, block chain and artificial intelligence have reshaped a whole new economic form, and digitalization is empowering businesses in many ways.

However, it cannot be ignored that many social problems in the development of digital economy are, to varying degrees, associated with the lack of social responsibility and alienation of platform enterprises. On the one hand, platform enterprises have many social responsibility deficiencies or alienation problems in their own business behavior, which bring serious harm to economic and social development. For example, take-out platforms have serious problems of consumer information security, and a large number of social responsibility deficiencies of illegal leakage and dumping of user information have occurred. On the other hand, the lack of management of irresponsible behaviors of bilateral users by platform-based enterprises leads to the supply or consumption behaviors of bilateral users relying on the
platform to have adverse effects on the economy and society, forming the second level of social responsibility deficiency of platform-based enterprises. Prominent examples include the vetting mechanism of online ordering platforms for “three no-go” take-out merchants \([5]\). The proactive social responsibility of platform enterprises is an essential element to maintain their long-term development momentum, which is the key to continuously promoting social and economic prosperity \([6]\). In addition, algorithmic hegemony, algorithmic discrimination, and algorithmic black box in the process of algorithmic services of platform enterprises are endless \([4]\). Algorithmic governance has also become a prominent research topic for socially responsible governance of platform companies in the digital era \([14]\).

Based on this, this paper focuses on three dimensions of CSR governance of platform enterprises in the digital context, and further analyzes the fourth new CSR governance dimension derived from algorithmic governance as a new CSR governance object. Specifically, this paper takes Meituan enterprise as the research object, analyzes the CSR development stage, and summarizes the CSR governance content from four dimensions to make up for the lack of attention to algorithm-embedded CSR governance issues in existing studies. It provides some guidance for the upgrading and development of emerging platform enterprises and the transformation of traditional enterprises, and lays the foundation for the ecological development of platform enterprises in the context of platform economy.

2 RESEARCH STATUS

Since Sheldon (1924) proposed the concept of CSR, the evolutionary logic of CSR has roughly gone through individual businessman-based social responsibility, stakeholder-responsive CSR, embedded social responsibility based on strategic competitive tool orientation, and platform value co-creation and sharing social responsibility based on value co-creation and value sharing logic orientation. With the gradual deepening of the penetration of "Internet+" into the economy and society, the traditional CSR content dimension is no longer applicable to platform enterprises, and the negative externalities of the platform economy on social responsibility seriously hinder the development of social and economic life \([11]\). Combining the development law of platform enterprise social responsibility and the characteristics of platform enterprises, Xiao Hongjun and Li Ping (2019) divide the content boundary of platform enterprise social responsibility into three levels: the first level is the social responsibility as an independently operating platform, and the implementation subject of its governance is the enterprise itself. Platform enterprises need to regulate their own behavior based on their own moral consciousness and the inner drive of self-awareness, and build relationships through the expression and response of each stakeholder's demands. The second level is the social responsibility as a platform for business operation. By exerting their own influence and control, platform enterprises play a role in the ecosystem, and all participants build a business ecosystem together based on the demand of value co-creation. The third level is the social responsibility as a social resource allocation platform. Based on their own characteristics, platform enterprises bring together various stakeholder subjects, bring into play the resource advantages of different subjects within the platform, build social-type platforms, integrate social resources and solve social problems.
With the accelerated use of the new round of digital technology, both traditional enterprises and platform enterprises have to some extent used products or services related to embedded algorithms, and algorithms have become the direct products and services of AI enterprises to achieve deep empowerment of traditional enterprises and reshape the productivity of the whole industry and even society \(^1\). Due to the highly opaque and uncontrollable nature of algorithms, stakeholders who apply algorithms or algorithm-embedded products and services need to be better informed about the possible and potential negative consequences of algorithmic decisions generating negative social issues such as algorithmic discrimination, algorithmic bias and algorithmic monopoly, which cannot be trusted by the outside world \(^2\). This paper assigns CSR as a platform for algorithm development and operation to the content dimension of platform CSR and re-examines how algorithm-embedded platform CSR is governed in the digital context. Figure 1 illustrates the governance mechanism of the four content dimensions of platform CSR.

3 CASE DESCRIPTION AND ANALYSIS

3.1 The Development History of Meituan and Key Events to Sort Out

Meituan was founded in 2010. As a life service platform company, Meituan has been diversifying its business to provide a series of services for the public, such as dining, entertainment, tourism, travel and shopping, which basically cover the basic areas of people's lives. Facing the rapid expansion of business and increasing social influence, Meituan keeps innovating and dynamically iterating its social responsibility strategy. In this paper, the key events of Meituan's social responsibility development are sorted out, as shown in the table 1.

<table>
<thead>
<tr>
<th>Time</th>
<th>Key events</th>
</tr>
</thead>
<tbody>
<tr>
<td>2017.06</td>
<td>Release of takeaway green ten.</td>
</tr>
<tr>
<td>2017.08</td>
<td>Green Mountain Program launched, industry's first to introduce no cutlery option.</td>
</tr>
<tr>
<td>2017.09</td>
<td>Establishment of Social Responsibility Committee, Establishment of Meituan</td>
</tr>
<tr>
<td>Year</td>
<td>Event</td>
</tr>
<tr>
<td>--------</td>
<td>----------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>2017.11</td>
<td>Launched &quot;Cell Guard&quot; applet.</td>
</tr>
<tr>
<td>2017.12</td>
<td>Take the lead in promoting takeaway confidence signatures nationwide.</td>
</tr>
<tr>
<td>2018.01</td>
<td>Held China's First Award Ceremony for Delivery Riders - &quot;New Urban Youth - 2018 Meituan Delivery Rider Award Ceremony&quot;.</td>
</tr>
<tr>
<td>2018.03</td>
<td>Launch of Rider Care's New Urban Youth Program.</td>
</tr>
<tr>
<td>2018.06</td>
<td>Open number protection function, Meituan public welfare platform officially launched.</td>
</tr>
<tr>
<td>2018.07</td>
<td>Launch of the 717 Knights Festival.</td>
</tr>
<tr>
<td>2018.08</td>
<td>Launch of Green Mountain Partner Program.</td>
</tr>
<tr>
<td>2018.10</td>
<td>Cooperation to explore the recycling of plastic boxes.</td>
</tr>
<tr>
<td>2019.10</td>
<td>Carrying out CSR ecosystem construction.</td>
</tr>
<tr>
<td>2020.01</td>
<td>Cooperation with several social entities to participate in the fight against the epidemic.</td>
</tr>
<tr>
<td>2020.09</td>
<td>People Magazine published the article &quot;Takeaway Riders, Stuck in the System.&quot;</td>
</tr>
<tr>
<td>2021.07</td>
<td>The General Administration of Market Regulation and other seven departments jointly issued &quot;on the implementation of the responsibility of the network catering platform to effectively safeguard the rights and interests of take-away food delivery personnel guidance.</td>
</tr>
<tr>
<td>2021.09</td>
<td>Announcing rules for rider delivery algorithms.</td>
</tr>
<tr>
<td>2021.10</td>
<td>Fined 3.442 billion yuan for &quot;choosing one over the other&quot;.</td>
</tr>
<tr>
<td>2022.03</td>
<td>50% commission reduction and targeted assistance for small and medium-sized merchants who have difficulties in operation.</td>
</tr>
</tbody>
</table>

### 3.2 Meituan Social Responsibility Philosophy

Building a social responsibility system plays an important role in Meituan's overall strategic planning. Meituan is committed to sharing its social responsibility philosophy across the entire ecosystem to promote the development and progress of all stakeholders and continuously create value for users, the industry and society. Meituan's social responsibility philosophy is practiced from three main aspects: users, industry and society.

1. **Users: Share the good life**
   
   Focusing on creating value for users, we are committed to meeting the different life needs of users, providing them with high-quality, full-scene consumer experience and contributing to people's high-quality life.

2. **Industry: Prosperous and beautiful industry together**
   
   Continuously strengthen scientific and technological innovation, promote the supply-side upgrade of the life service industry, and promote the high-quality development of the industry through various initiatives to help cultivate talents in the industry.

3. **Society: Building a better society together**
   
   We actively explore the advantages of using technology and platform to help solve more social problems, create greater social value, and work together with related parties to build a sustainable and beautiful society.
3.3 Two Major Algorithms of Meituan

In the operation mechanism of Meituan's algorithm, the algorithm strives to achieve the optimal solution for the needs of riders, users and merchants together, which ultimately leads to the sustainable development of the platform.

Distribution algorithm. Fast delivery is one of Meituan Takeaway's core competencies. Behind the surge in corporate net profit is Meituan's continuously improved and optimized algorithm. Specifically, Meituan Takeaway uses big data to perform calculations that can provide consumers with the most optimized delivery path in the shortest possible time. In emergency situations, the machine has a strong learning capability and can adjust its strategy through various algorithms to make efficient millisecond dispatching of orders. Meituan Delivery continuously improves delivery efficiency and enhances market competitiveness through accurate delivery algorithms.

Recommendation algorithm. With the rapid expansion of user group size, Meituan has accumulated diversified user behavior information. In the new development stage, the double-wheel mechanism of data and algorithm can empower the platform technology innovation [2]. Data is the basis of algorithms and models, and these data provide essential conditions for the application and optimization of recommendation systems. After different types of user data are collected and classified, they are ultimately used in algorithms and models. Simple data is just byte stacking, while algorithms can grasp the principles and laws of the data through the calculation of the data, and bring the value of the data to the maximum. The core of collaborative filtering is to find and recommend "similar interests", which simply means to recommend information that users may be interested in with the help of their common hobbies and experiences to meet their needs for personalized services.

4 CASE STUDY

Compared with traditional enterprises in a general sense, platform enterprises have stronger public and social attributes, and the business ecosystem they build is more complex, so their social responsibility presents governance complexity [10], subject multiplicity [9], mixed content [6], and service sociality [7]. In terms of the function of the role of platform enterprises in society, they are both independent operating subjects with the characteristics of general commercial enterprises and commercial operation platforms linking bilateral or multilateral users and members of other ecological niches, as well as social resource allocation platforms with the potential to gather resources of a larger range of social subjects and algorithm development and operation platforms spawned by the digital economy.

4.1 Performance Role Analysis

This paper deepens the social responsibility ecosystem based on Meituan's fourfold role positioning as an independent operating entity, commercial operation platform, social resource allocation platform and algorithm development and operation platform, and identifies four levels of the social responsibility ecosystem and the corresponding social responsibility levels. First, it is the social responsibility governance as an independent operating enterprise subject, focusing on the awareness of responsibility to the platform stakeholders, of which the
The protection of consumer rights and interests can be a key initiative of Meituan. The 2021 CSR Report states that the main business scenarios under the Meituan platform have achieved user number privacy protection, and also established a strict data security management system to ensure that only authorized personnel can access internal information.

The second is the social responsibility governance as a commercial operation platform, focusing on the governance of merchants' social responsibility and the cultivation of other stakeholders to participate in the common governance awareness of social responsibility. Meituan has developed and built the "electronic file system for merchants in the network", and through the three links of "entry audit, registration in the network, and exit tracking", it manages the whole life cycle of the merchants in the network to avoid the inflow of black workshops into the market. As a "platform governor", the important goal of Meituan's social responsibility governance is to restrain and regulate the socially responsible behavior of bilateral users and maintain the standardized operation of the platform.

Third, socially responsible governance as a social resource allocation platform is reflected in improving the degree of social participation and enhancing the platform's ability to integrate social resources and solve social problems. In 2017, Meituan launched its first environmental initiative in the takeaway industry, the Green Mountain Plan, to explore the green transformation of the takeaway industry in cooperation with ecological parties. In 2021, Meituan will upgrade the environmental strategy of the Green Mountain Plan, not only increasing investment in green packaging for takeaway restaurants, but also emphasizing the deep integration of green and low-carbon concepts into all businesses and products, and widely uniting other platform stakeholders to play the linkage advantage, and jointly promote the industry's environmental protection process.

The fourth is socially responsible governance as an algorithm development and operation platform, which is reflected in the promotion of transparent and quality operation of the platform. The algorithm development platform is the fourth role position of Meituan Takeaway, an intelligent platform that owns the development and operation of algorithm programs. As the social responsibility ecology of the algorithm development platform, the ecological partners of which join the algorithm as the subject of governance. For the Meituan takeaway rider delivery program algorithm black box problem, on the one hand, the platform enterprise itself algorithm transparency governance, Meituan enterprise itself through the algorithm source and governance system system disclosure; on the other hand, the joint various ecological partners, its governance responsibility empowerment, around the algorithm black box, discrimination, social responsibility issues, the establishment of supervision and management and public opinion guidance measures to maximize the transparency of the algorithm.

4.2 Algorithm-based Embedded Socially Responsible Governance Model

4.2.1 Rationalization of Disclosure

An algorithm can be understood as a specific mathematical computational model that implements a series of coding, procedural and logical rules for transforming data into the corresponding output results. A complete algorithmic system can be self-iterating and optimized based on a data set, a process also known as machine learning. Its formation and
operation consists of four main stages: "input-learning-output-application". The input stage is to collect and classify the corresponding data under the condition of defining the basic problem, to clean and categorize the data, and on the basis of this, the learning stage is to continuously train the data and the performance of the model, to output the series of results of the model to evaluate the accuracy of the output results, to ensure that the output results of the model can reach the established expected calculation and decision goals, and finally to achieve the optimal results. The algorithmic model is applied to the actual situation.

According to the research status of algorithm disclosure by domestic and foreign scholars, rationalized disclosure is mainly carried out at two levels: the source of algorithms and the consequences of operation. From the source level of algorithms, we mainly focus on the "input" and "learning" aspects. Currently, algorithm disclosure can be divided into three levels: primary disclosure, specific transparency and full disclosure. Our platform companies always maintain the principle of algorithm transparency and explainability, and emphasize the disclosure of the existence of algorithm automation decision, which can be regarded as the primary disclosure of algorithm disclosure. With the increase of requirements for transparency of algorithmic automated decision making, platform enterprises begin to disclose the model, operation logic, principles, rules, and specific algorithmic parameters of specific algorithmic automated decision making. It can be regarded as the second disclosure level of algorithm specific transparency. Besides, the proposal of disclosing the source code of algorithms in the United States, which requires the providers of algorithmic automated decision making services to disclose the source code and other relevant contents about algorithmic automated decision making, can be regarded as the full disclosure of algorithms. At present, the U.S. has not had any relevant disclosure behavior about the source code of the algorithm, and its practice only stays at the second disclosure level. From the perspective of the operational consequences of the algorithm, the "output" and "application" aspects are the focus of implementation. The uncontrollability and uncertainty of algorithm automation decisions dictate the need to strengthen the disclosure of the possible consequences of algorithm operation. At the internal level, we need to build a system of algorithm governance, such as algorithm disclosure system, algorithm transparency management system, and algorithm impact assessment system. In November 2021, Meituan disclosed the algorithm rules of "order allocation" at the source of the algorithm, introducing the logic of matching orders and riders. In May 2022, Meituan optimized its takeaway algorithm, adding a new response to "abnormal scenarios". Meituan mainly discloses the operation logic, purpose, principle and final effect of the algorithm's automated decision making. At the same time, Meituan expresses the operation rules of "order allocation algorithm" in easy-to-understand language, and illustrates and justifies the application of the algorithm in the form of icons, diagrams and examples, which is a powerful attempt of detailed disclosure of the algorithm in China.

4.2.2 Responsibility Empowerment

Under the dual attribute of platform, based on the platform business ecosystem theory, the social responsibility governance of platform enterprises is different from the traditional unilateral internal social responsibility management of enterprises or linearized governance of supply chain members, and platform enterprises form two types of ecological niches based on the unique platform business ecosystem structure (Moore J F, 1993). The first type of main ecological niche is based on platform enterprises as core members; the second type of
extended ecological niche, whose main covered subjects include other competing platform enterprises, government, social organizations and social public. This type of ecological niche has multiple roles of forming guidance, supervision and governance for the algorithmic behavior of the main ecological niche members. In this paper, we define responsibility empowerment as the main ecological niche members carry out traction and empower the governance responsibility of the expanded ecological niche subjects in the social responsibility ecosystem. It is mainly divided into two aspects: government regulation and social supervision.

Government regulation. The government's scientific regulation of algorithms, the basic principles of inclusive and prudent regulation and graded and classified regulation, the construction of an agile governance-based dynamic and interactive regulatory system, and the promotion of more refined and targeted legislation and law enforcement while adapting to the characteristics of algorithm governance and the development needs of algorithm innovation. In recent years, China has attached great importance to legislation in related fields. The introduction of the Data Security Law, the Personal Information Protection Law, the Network Security Law and other related laws proves that China attaches great importance to ethics and governance in algorithm-related innovation. In the era of digital economy, the mastery of big data is the mastery of great power. As platform companies use big data, they are pushing labor alienation to the extreme. Takeaway riders lose control of their work and have to rely entirely on the platform's algorithm to distribute orders on an ad hoc basis. In a data- and algorithm-driven platform environment, continuous, forward-looking regulation is essential, and scientific government regulation of algorithms is a prerequisite to ensure the coexistence of market order and vitality, safety and development. In July 2021, the State Administration of Market Supervision and other seven departments jointly issued the "Guidance on the implementation of network catering platform responsibility to effectively safeguard the rights and interests of take-away food delivery workers", to protect the legitimate rights and interests of take-away food delivery workers to put forward a full range of requirements.

Social oversight. In addition to state power, news agencies, social organizations and individual citizens form an algorithmic governance community. All-round and deep supervision of practices and behaviors in algorithm applications that violate citizens' rights, affect market competition, social order, public safety, and national security can better achieve algorithmic good governance. News media supervision has promoted the iterative update of algorithms in some specific scenarios such as news recommendation and delivery time estimation. After an article titled "Delivery riders, trapped in the system" was widely disseminated through self-media in September 2020, the working status of online delivery workers under the control of big data algorithms gained the attention of all sectors of society. The community's concern for monitoring the algorithm-focused incident has generated extensive social discussion and has driven a positive response from lawmakers. Individual citizens often participate in algorithm governance through complaints, reporting, and filing lawsuits. At present, the overall system of citizens' digital rights in China is still relatively vague, and the protection of digital rights still needs to be improved. In the future, platform enterprises can establish to accept social supervision, set up convenient complaint reporting entrances, receive and handle public complaints and reports in a timely manner; establish user complaint channels and systems, handle user complaints in a standardized manner and provide feedback on complaints in a timely manner. These provisions are not only convenient for citizens to defend their rights, but also can form a social public supervision mechanism for platform enterprises.
5 CONCLUSION

Synthesizing past scholars' studies, there has been a great deal of research in academia on the new economic form shaped by digital change, the innovative changes of micro enterprises, and the continuous improvement of production efficiency. Meanwhile, elements related to CSR governance of platform enterprises have also attracted extensive attention [3]. However, few studies have systematically analyzed the algorithm-embedded platform CSR governance. In this paper, the new content dimension of algorithm development and operation platform is given to the platform enterprises, and a systematic analysis and innovative framework for the new CSR in the digital context is sorted out.

This paper integrates the social responsibility of platform companies in four dimensions: "social responsibility as an independent operating entity", "social responsibility as a commercial operation platform", "social responsibility as a social resource allocation platform", and "social responsibility as an algorithm development and operation platform". This paper integrates CSR governance in four dimensions: "CSR as a platform for algorithm development and operation". In this paper, the algorithm governance under CSR governance is grouped into two dimensions, including rationalization of disclosure and responsibility empowerment. On the one hand, the disclosure of data sources, data collection standards and algorithm learning training standards (data analysis and processing process) of algorithms is strengthened at the source level of algorithms, and on the other hand, the disclosure system of potential impact and risk assessment of algorithm decisions is built, and finally, the responsibility disclosure system of algorithms, algorithm transparency management system and responsibility assessment of algorithm impact are built at the internal level of enterprises. In the end, the algorithm governance system such as the algorithm responsibility disclosure system, the algorithm transparency management system, and the algorithm impact responsibility assessment system are built at the internal level. From the perspective of responsibility empowerment, platform companies empower the governance responsibilities of each stakeholder in the social responsibility ecosystem, so that the rights and obligations of algorithmic governance can be specified and clarified. Through government guidance and social supervision, algorithm governance is promoted to create a "good" algorithm ecology.

REFERENCES

“Avoid a Crisis” or “Out of the Crisis”: A Study Based on the Impact of Manager Competence on Organizational Resilience of Small and Medium-Sized Enterprises

Mengsu Xu1,a*, Zilin Liu2,b, Yiwei Wang3

*a1359008965@qq.com, xms0623@126.com

Yifu Building, East District, Qiqihar University, jianhua district Culture Street, Qiqihar City, Heilongjiang Province, China1

Shandong University of Finance and Economics Jinan City, Shandong Province, China2

Anhui University of Finance and Economics, Bengbu City, Anhui Province, China3

Abstract: The frequent public crises in recent years have revealed that small and medium-sized enterprises lack crisis management capabilities and organizational resilience. Under the stimulus of external crisis, the managers' ability of small and medium-sized enterprises has a great impact on the risk-taking and resilience level of enterprises. Taking small and medium-sized enterprises in Shanghai and Shenzhen A-shares as samples, this paper empirically analyzes the relationship among managers' ability, corporate risk-taking level and organizational resilience. The results show that the stronger the ability of managers, the higher the resilience level of small and medium-sized enterprises, and the higher the risk-taking level of enterprises. In addition, the enterprise's risk-taking level plays a part of intermediary role in the relationship between managers' ability and the organizational resilience of small and medium-sized enterprises; The digital transformation is positively adjusting the relationship between managers' ability and the level of corporate risk-taking. The research results enrich the relevant literature and provide a certain reference for the selection of managers and digital transformation of enterprises.

Keywords: Manager's Ability, Organizational Resilience, Enterprise Risk-taking Level, Digital Transformation, Small and Medium-sized Enterprises.

1 INTRODUCTION

Connecting small and medium-sized enterprises with thousands of households is an important force to promote innovation, promote employment and improve people's livelihood. However, in recent years, public crises have occurred frequently, and the global economy has become increasingly complex. Small and medium-sized enterprises are facing various unknown risks, financing problems, internal management problems, etc. Because of their small scale, insufficient funds and poor liquidity, their competitiveness is weak in the face of sudden market fluctuations and policy changes, and it is urgent to improve their ability to cope with crises and improve their organizational resilience. Resilience is generally regarded as an ability to effectively deal with crises, avoid injuries and develop further. Especially for high-
impact and low-probability events, it can realize early warning, form redundant supply, reduce vulnerability and increase flexibility. In the fierce market competition environment, how can small and medium-sized enterprises ensure their sustainable development? In addition, under the background of China's vigorous development of digital economy, the digital transformation of enterprises is an inevitable trend of economic development. Will the digital transformation of enterprises affect the risk-taking level of enterprises? In order to answer this question, this paper empirically tests the relationship between managers' ability and the resilience of small and medium-sized organizations, taking China's A-share listed small and medium-sized companies from 2010 to 2021 as samples, and further explores the intermediary role of risk-taking and the regulatory role of digital transformation.

The possible research contributions of this paper are as follows: First, focusing on small and medium-sized enterprises, this paper studies the influence of their management ability on their organizational resilience, enriches the research content on the sustainable development of small and medium-sized enterprises, and is of great significance to the realization of specialized and new development of small and medium-sized enterprises. Secondly, the application of dynamic capability theory is extended and the antecedents of organizational toughness are tested. The possible mechanism of the formation of organizational toughness is clarified theoretically, which is helpful to expand the mechanism research on the influence of managers' ability. Thirdly, the digital transformation is brought into the research framework, and the impact of the degree of enterprise digital transformation on the managers' ability and the risk-taking level of small and medium-sized enterprises is explored, which enriches the research content of digital transformation and provides countermeasures and reference for promoting the digital transformation of small and medium-sized enterprises.

2 LITERATURE REVIEW AND RESEARCH HYPOTHESIS

2.1 Managerial Ability and Organizational Resilience

The manager's ability is embodied in many aspects, including the manager's knowledge, experience and values. It reflects the manager's professional ability to deal with complex issues such as opportunity identification, risk taking and resource integration. It plays a very important role in small and medium-sized enterprises to obtain sustainable competitive advantages. Based on the high-level theory, when an enterprise is in crisis, it is necessary to have a hero who can save the crisis, i.e. a manager with superior ability. They often show extraordinary learning ability and calling spirit, that is, they can keenly perceive the changes in the external environment, identify the risk factors that may cause disasters to the enterprise, activate the collective wisdom and initiative of the employees of the company, and implant resilience genes for the sustainable development of small and medium-sized enterprises. In small and medium-sized enterprises, the manager is the key figure in the enterprise organization, which always affects the subordinate's behavior and team atmosphere. The leader's ability to think strategically during a crisis determines whether an organization can survive for a long time. He is at the top of the pyramid in the organization's power position and is considered as an important trigger for resilience. When there are challenges, crises and other stimulus factors in the environment, the employees of the enterprise often do not have such resilience consciousness. They need the stimulation of managers with crisis
consciousness to trigger the collective psychology of the team to deal with together, and then produce team resilience, which is also called "peer effect"[10]. Therefore, the key to the development of organizational resilience in small and medium-sized enterprises is to cultivate the ability of managers. Based on this, hypothesis 1 is proposed:

H1: Managers' ability has a positive impact on the organizational resilience of SMEs.

2.2 Manager's Ability and Enterprise's Risk-taking Level

Corporate risk-taking reflects the overall investment strategy of the enterprise and the risk preference of the managers, while the managers' ability reflects the managers' cognitive level and ability to handle affairs[3]. According to the "risk-taking hypothesis" of the butler's theory in modern management, the more capable the manager is, the more capable he is of taking risks. First of all, high-caliber managers often have a different vision from ordinary people. They can accurately judge customers' needs in the rapidly changing market environment, and can also find investment opportunities, assess the value of potential investment opportunities, and effectively control risks when improving the efficiency of enterprises' investment projects. As a result, the level of enterprises' risk-taking will be higher. Secondly, managers with strong capabilities generally have strong social resources and relationship networks, which have an important impact on the establishment of a stable and sustainable trading model. They can enhance the timeliness and stability of resource supply in business activities, optimize and integrate the allocation of enterprise resources, reduce the risks in the decision-making process, and provide resource guarantee for high-risk projects. In addition, the more capable managers are, the better they will be able to prepare a complete risk response plan in advance. Even if a crisis occurs, they will be able to calmly respond with professional skills, minimize losses and realize the sustainable development of the enterprise. Therefore, based on the above analysis, the following assumption 2 is proposed:

H2: The ability of managers has a positive impact on the level of corporate risk-taking.

2.3 The Intermediary Role of Enterprise Risk-taking Level

Most of the listed companies of small and medium-sized enterprises are in the growth stage, and they are high-growth and small-scale enterprises. Financing is relatively difficult and has a great demand for financing. As the person in charge of enterprise management and investment decision-making, managers' own management ability will naturally affect the final investment and decision-making of enterprises, that is, the level of enterprise risk-taking[2]. In addition, some strategic plans made by the company are usually accompanied by potential risks, such as mergers and acquisitions, diversified businesses, new product research and development, etc., which are often related to the management ability of small and medium-sized enterprises. According to the dynamic capability theory, enterprises can reconstruct internal resources according to environmental changes, quickly integrate existing resources, and realize enterprise development. Organizational toughness is an important dynamic capability of an enterprise. The stronger the ability of managers, the higher the risk-taking level of enterprises will be, not only can they control the losses in the process of strategy implementation, but more importantly, they can use their own risk management ability to gain the advantage of continuous competition, that is, according to environmental changes, quickly integrate resources and cope with risks, that is, the organizational resilience of enterprises. In small and
medium-sized enterprises, the two rights of the company are basically in the hands of the same leader, and the risk-taking level of the enterprise mostly depends on the decision of the manager. Keri Ultrasonic Electronics Co., Ltd., a small and medium-sized enterprise in Foshan, Guangdong Province (hereinafter referred to as Keri Ultrasonic), was founded at the beginning, with many homogeneous enterprises, fierce market and low profit rate. In the case of continuous losses, most enterprises are unwilling to invest time and net profit to improve quality, but choose to reduce prices and thin profits to avoid risks. Ye Weizhong, the leader of Keri Ultrasound, has always insisted on investing in research and development of new products. Under the unified leadership of Ye Weizhong, he has continuously optimized product quality, appearance design and cost, constantly made progress against risks, firmly grasped the core technology, and maintained enough toughness and endurance again and again in the crisis, making it invisible to produce ultrasonic atomizing tablets in the high-end atomizing market. Therefore, the higher the ability of managers, the more they can enhance the risk-taking level of enterprises and further enhance the resilience of small and medium-sized enterprises. Based on the above analysis, this paper puts forward hypothesis 3:

H3: The level of enterprise risk-taking has an intermediary effect in the process of managers' ability affecting the organizational resilience of SMEs.

2.4 The Regulatory Role of Digital Transformation

Digital transformation is defined as the transformation of information technology\(^4\), which generally involves changes in business processes, operating procedures\(^3\) and organizational capabilities\(^8\) as well as ways to enter new markets or exit current markets. Digital transformation is characterized by the application of artificial intelligence, blockchain, cloud computing, big data and other underlying technologies in the organizational structure and management mode of enterprises, which may affect the willingness to take risks and financial situation of enterprises by influencing management behavior, reshaping the internal governance system of enterprises and improving the external environmental constraints of enterprises. Therefore, digital transformation can alleviate the agency contradiction between owners and management, reduce agency costs, thereby improving the level of enterprise risk, and provide a good environment for the improvement of enterprise risk. In addition, the digital infrastructure that enterprises rely on for digitalization can improve managers' ability, help managers to capture business information more comprehensively, and ensure that the implementation of business processes is more stable and reliable. In the digital environment of enterprises, the control constraints of managers and ordinary employees will not vary from person to person, and the internal information communication of enterprises will be more transparent and smooth, which will further enhance the effectiveness of control environment, control activities, information communication and supervision. Therefore, the digitalization of enterprises has a regulating effect on the relationship between managers' ability and enterprise risk-taking level\(^9\). Based on this, this paper puts forward hypothesis 4:

H4: Digital transformation has a positive regulatory effect on the relationship between managers' ability and the risk-taking level of SMEs.

The conceptual model studied in this paper is shown in Figure 1.
3 VARIABLE INDICATORS AND EMPIRICAL MODELS

3.1 Data Sources

The data are obtained from WIND and CSMAR Database. The sampling time is from 2010 to 2021 and the sample size is 15801. The companies with serious missing data samples, asset-liability ratio greater than 100% and financial and insurance industries are excluded. Finally, the continuous variables are trimmed by 1% to 99%.

3.2 Variable Selection

3.2.1 Explained Variable

Organizational resilience. Learn from Lv [6] to test organizational resilience from two dimensions: long-term growth and financial fluctuation. Long-term growth is measured by the accumulation of three-year net sales growth, and financial fluctuation is measured by the return on stocks. For the annual stock return, the standard deviation of the monthly stock return is calculated first, and then the annual stock return is calculated according to the monthly stock return. Finally, the Organizational resilience is comprehensively calculated by entropy method.

3.2.2 Explanatory Variable

Manager's ability. This paper uses the DEA-TOBIT model to measure managers' ability (MA) by referring to the research of Demerjian[1]. The specific steps of this method are as follows: In the first stage, the data DEA efficiency model is used to measure the productivity of enterprises by industry and year. Taking Revenue as the output variable, operating cost (OCC), sales expenses and management expenses (SME), net fixed assets (NFA), net intangible assets (NIA), goodwill (BR) and R&D expenditure (RD) as input variables, the following model is established.

\[
MAX(\delta) = \frac{Revenue}{\alpha_0 OCC + \alpha_1 SEM + \alpha_2 NFA + \alpha_3 NIA + \alpha_4 BR + \alpha_5 RD} \tag{1}
\]
In the second stage, the manager's ability is estimated. The efficiency calculated by the above formula includes both the efficiency of the manager's ability creation and the efficiency of the company's idiosyncratic creation. Therefore, in order to eliminate the influence of corporate level on efficiency, this paper selects the following six factors that may affect the production efficiency of the enterprise: corporate size (CS), market share (MS), free cash flow (FCF), listing life (LY), diversified operation (DO) and overseas operation (OB) for Tobit regression, and the residuals obtained from the regression can be used to measure managers' ability (MA). The specific calculation model is as follows:

$$\delta = \beta_1 CS + \beta_2 MS + \beta_3 FCF + \beta_4 LY + \beta_5 DO + \beta_6 OB + \varepsilon$$

(2)

### 3.2.3 Intermediary Variable

Enterprise Risk Undertaking (RiskT). This paper uses the earnings volatility during the observation period of the enterprise to measure the enterprise risk-taking level (RiskT), and uses the practice of Yu Minggui [12] and others for reference. First, calculate ROAi, which is the ratio of the enterprise's profit before tax, interest, depreciation and amortisation for Year I to the total assets at the end of the year. Second, the volatility is calculated by first adjusting the industry average adopted by the enterprise every year, and then calculating the standard deviation adjusted by the industry in each observation period. The risk exposure level (RiskT) of the enterprise is as follows:

$$RiskT_{it} = \frac{1}{T-1} \sum_{t=1}^{T} (AdjROA_{it} - \frac{1}{T} \sum_{t=1}^{T} AdjROA_{it})^2$$

(3)

$$AdjROA_{it} = \frac{EBIT_{it}}{Asset_{it}} \cdot \frac{1}{X} \sum_{k=1}^{X} \frac{EBIT_{ik}}{Asset_{ik}}$$

(4)

i represents the enterprise and t represents the year of the observation period.

### 3.2.4 Regulatory Variables

Digital transformation. Drawing on the experience of Wu Fei [11] in their research on digital transformation, they selected five key words about digital as feature words, including artificial intelligence technology, big data technology, cloud computing technology, blockchain technology and digital technology application. Using Python statistical annual report in the frequency of the occurrence of keywords, sum up to digital transformation of the total index, and then take the natural log measurement.

### 3.2.5 Control Variables

Drawing lessons from previous research, the control variables are Size and Age of enterprise, debt-to-asset ratio, executive compensation, government subsidy, return on total assets, dual, independent director ratio, cash flow, shareholder holding ratio, sales growth rate, year and industry.
3.3 Model Design

Construct the panel regression benchmark model in the following form:

\[ \text{Org}_{i,t} = \alpha_0 + \alpha_1 \text{MA} \cdot \text{Score}_{i,t} + \sum \alpha_2 \text{Control}_{i} + \sum \text{Ind} + \sum \text{Year} + \epsilon_{i,t} \]  \hspace{1cm} (5)

Model (1) is used to verify hypothesis 1. \( \alpha_0 \) is the intercept term, where \( \alpha_i \) is the regression coefficient and \( \epsilon \) is the random error term. If the coefficient of MA is significantly positive, it indicates that managers' ability can promote organizational resilience.

In order to further investigate the intermediary role of enterprise risk-taking level, according to the commonly used three-step method of intermediary role test, the following extended model is constructed:

\[ \text{Org}_{i,t} = \alpha_0 + \alpha_1 \text{MA} \cdot \text{Score}_{i,t} + \sum \alpha_2 \text{Control}_{i} + \sum \text{Ind} + \sum \text{Year} + \epsilon_{i,t} \]  \hspace{1cm} (6)

\[ \text{RiskT}_{i,t} = \beta_0 + \beta_1 \text{MA} \cdot \text{Score}_{i,t} + \sum \alpha_3 \text{Control}_{i} + \sum \text{Ind} + \sum \text{Year} + \epsilon_{i,t} \]  \hspace{1cm} (7)

\[ \text{Org}_{i,t} = \gamma_0 + \gamma_1 \text{MA} \cdot \text{Score}_{i,t} + \gamma_2 \text{RiskT} + \sum \alpha_4 \text{Control}_{i} + \sum \text{Ind} + \sum \text{Year} + \epsilon_{i,t} \]  \hspace{1cm} (8)

Equations (6)-(8) focus on the parameters \( \alpha_1, \beta_1, \gamma_1 \) and \( \gamma_2 \). If \( \beta_1 \) is significant, it indicates that the manager's ability will affect the risk-taking level of the enterprise. On this basis, if \( \gamma_1 \) and \( \gamma_2 \) are both significant and the absolute value of \( \gamma_1 \) is less than the absolute value of \( \alpha_1 \), then it is considered that the enterprise's risk-taking plays a part of intermediary role between managers' ability and organizational resilience; If \( \gamma_2 \) is significant, but \( \gamma_1 \) is not significant, it indicates that risk-taking plays a complete mediating role between managers' ability and organizational resilience.

In order to examine the moderating effect of digital transformation on managers' ability to bear corporate risks, the following expansion model was constructed:

\[ \text{Org}_{i,t} = \alpha_0 + \alpha_1 \cdot \text{MA} + \alpha_2 \cdot \text{DCG} + \alpha_3 \cdot \text{MA} \cdot \text{DCG} + \sum \alpha_4 \text{Control}_{i} + \sum \text{Ind} + \sum \text{Year} + \epsilon_{i,t} \]  \hspace{1cm} (9)

In formula (9), the significant regularity of \( \alpha_3 \) coefficient indicates that the digital transformation has a positive moderating effect on the manager's ability and the level of enterprise risk commitment; otherwise, it has a negative moderating effect.

4 EMPIRICAL ANALYSIS

4.1 Descriptive Statistics

Table 1 makes descriptive statistics on the main variables. It can be seen from the table that there is a big gap between the resilience level of SMEs and the ability of managers. The
average risk-taking of enterprises is 0.027, which shows that the overall risk-taking degree of small and medium-sized enterprises is not very high. In terms of digital transformation, the average value of control variables is close to the median value as a whole, so it shows that the sample does not have a big deviation.

**Table 1** Descriptive Statistics

<table>
<thead>
<tr>
<th>Variables</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Median</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Org</td>
<td>15801</td>
<td>4.041</td>
<td>1.744</td>
<td>0.65</td>
<td>3.764</td>
<td>96.617</td>
</tr>
<tr>
<td>MA</td>
<td>15801</td>
<td>-0.04</td>
<td>0.171</td>
<td>-0.6</td>
<td>-0.03</td>
<td>0.545</td>
</tr>
<tr>
<td>RiskT</td>
<td>15801</td>
<td>0.027</td>
<td>0.026</td>
<td>0</td>
<td>0.019</td>
<td>0.118</td>
</tr>
<tr>
<td>DCG</td>
<td>15801</td>
<td>2.075</td>
<td>1.118</td>
<td>0.69</td>
<td>1.946</td>
<td>4.356</td>
</tr>
<tr>
<td>Age</td>
<td>15801</td>
<td>20.01</td>
<td>6.075</td>
<td>2</td>
<td>20</td>
<td>67</td>
</tr>
<tr>
<td>Lev</td>
<td>15801</td>
<td>44.05</td>
<td>20.19</td>
<td>1.1</td>
<td>43.71</td>
<td>195.66</td>
</tr>
<tr>
<td>Pay</td>
<td>15801</td>
<td>5.365</td>
<td>0.738</td>
<td>-0.17</td>
<td>5.327</td>
<td>9.081</td>
</tr>
<tr>
<td>Gov</td>
<td>15801</td>
<td>16.5</td>
<td>1.587</td>
<td>0</td>
<td>16.5</td>
<td>22.735</td>
</tr>
<tr>
<td>ROA</td>
<td>15801</td>
<td>4.05</td>
<td>7.626</td>
<td>-93</td>
<td>4.006</td>
<td>96.864</td>
</tr>
<tr>
<td>Dual</td>
<td>15801</td>
<td>0.29</td>
<td>0.454</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Indep</td>
<td>15801</td>
<td>37.68</td>
<td>5.588</td>
<td>14.3</td>
<td>36.36</td>
<td>80</td>
</tr>
<tr>
<td>CF</td>
<td>15801</td>
<td>1.060</td>
<td>7.240</td>
<td>-4.34</td>
<td>1.58</td>
<td>3.667</td>
</tr>
<tr>
<td>INS</td>
<td>15801</td>
<td>39.15</td>
<td>23.98</td>
<td>0</td>
<td>39.73</td>
<td>326.73</td>
</tr>
<tr>
<td>Top1</td>
<td>15801</td>
<td>33.69</td>
<td>14.99</td>
<td>2.43</td>
<td>31.2</td>
<td>89.09</td>
</tr>
<tr>
<td>Grow</td>
<td>15801</td>
<td>0.515</td>
<td>17.3</td>
<td>-0.95</td>
<td>0.126</td>
<td>1878.4</td>
</tr>
</tbody>
</table>

4.2 Correlation Analysis

Table 2 shows the Pearson correlation coefficient test results for the study variables. The correlation coefficient between each variable is less than 0.5, and the VIF value of each variable is far less than 2. Therefore, there is no multicollinearity problem in the regression model.

**Table 2** Analysis of Regression Results

<table>
<thead>
<tr>
<th>Variables</th>
<th>Org</th>
<th>MA</th>
<th>RiskT</th>
<th>DCG</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Org</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1.19</td>
</tr>
<tr>
<td>MA</td>
<td>0.039***</td>
<td>1</td>
<td></td>
<td></td>
<td>1.16</td>
</tr>
<tr>
<td>RiskT</td>
<td>0.077***</td>
<td>0.077***</td>
<td>1</td>
<td></td>
<td>1.07</td>
</tr>
<tr>
<td>DCG</td>
<td>0.023***</td>
<td>0.054***</td>
<td>0.041***</td>
<td>1</td>
<td>1.08</td>
</tr>
</tbody>
</table>

Note: Due to space reasons, only the correlation results of main variables are shown.

4.3 Analysis of Empirical Results

4.3.1 Management Ability and Organizational Resilience

Through the White test, it is found that the results reject the original assumption that the sample may have heteroscedasticity. Referring to relevant literature, this paper uses the generalized least squares (GLS) to solve the problem of heteroscedasticity. The regression results are shown in Table 3. Model 1 is the regression result of manager's ability to enterprise's organizational resilience. The MA coefficient in Model 1 is 0.215 and significant at the level of 1%. The sign is positive, i.e. there is a positive relationship between manager's ability and enterprise's organizational resilience, indicating that manager's ability is very
important. The stronger manager's ability, the stronger enterprise's organizational resilience. Hypothesis 1 is verified.

4.3.2 The Ability of Managers and the Level of Enterprise Risk-taking

Model 2 is the regression result of managers' ability to undertake risks to the enterprise. Among them, the coefficient of management ability (MA) is 0.015 and is significantly positive at 1%, indicating that with the improvement of management ability, the enterprise's risk-taking level will be higher. Hypothesis 2 is verified. When facing crises and challenges, high-caliber managers have better opportunities to explore, identify risks and integrate resources. With rich corporate governance experience, they can keenly detect the information contained in environmental changes, identify potential risks in innovative projects, quickly adjust strategies, avoid strategic risks and improve the risk-taking level of small and medium-sized enterprises.

Table 3 Analysis of Regression Results

<table>
<thead>
<tr>
<th>Variables</th>
<th>(1)Org</th>
<th>(2)RiskT</th>
<th>(3)Org</th>
<th>(4)RiskT</th>
</tr>
</thead>
<tbody>
<tr>
<td>MA</td>
<td>0.215*** (13.55)</td>
<td>0.015*** (29.96)</td>
<td>0.139*** (7.53)</td>
<td>0.008*** (7.69)</td>
</tr>
<tr>
<td>RiskT</td>
<td>5.057*** (58.88)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DCG</td>
<td>0.001 (1.62)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MA×DCG</td>
<td>0.031*** (7.18)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Controls</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>cons</td>
<td>4.667*** (115.19)</td>
<td>0.051*** (43.92)</td>
<td>4.431*** (86.03)</td>
<td>0.051*** (39.61)</td>
</tr>
<tr>
<td>N</td>
<td>15801</td>
<td>15801</td>
<td>15801</td>
<td>15801</td>
</tr>
<tr>
<td>Chi-squared</td>
<td>93496.578***</td>
<td>28460.347***</td>
<td>188635.710***</td>
<td>65614.445***</td>
</tr>
</tbody>
</table>

Note: T statistics are in brackets, and * *, * * and * represent the significance levels of 1%, 5% and 10% respectively, as above.

4.3.3 The Intermediary Role of Risk taking

Model 3 is the regression result after adding the intermediate variable of risk taking (RiskT). It can be seen that the coefficient of manager ability (MA) is 0.139, which is significantly positive at the level of 1%, and the coefficient of risk taking (RiskT) is 5.057, which is also significantly positive at the level of 1%. According to the test method of mediating effect, as the managerial competence (MA) of model 1 and model 2 is positively significant to organizational resilience and enterprise risk-taking level at 1%, and the managerial competence (MA) and risk-taking (RiskT) of model 3 are also positively significant to organizational resilience at 1%, therefore, The test proves that risk-taking plays a part in mediating the effect of managers' ability on the organizational resilience of an enterprise, i.e. the higher the managers' ability, the higher the risk-taking level of an enterprise can be, thus further promoting the organizational resilience of small and medium-sized enterprises. Hypothesis 3 is verified. The sustainable development of small and medium-sized enterprises cannot be achieved without the performance of the enterprise's organizational resilience. The stronger the manager's ability, the more flexible he will be in dealing with the risks and
challenges of the uncertain market. As a result, the enterprise's risk-taking level will also be improved, and the resilience of small and medium-sized enterprises will also be stronger.

### 4.3.4 The Regulating Function of Digitalization

Model 4 adds the interaction item (MA×DCG) between managers' ability and digital transformation to test the moderating effect of digital transformation (DCG). The regression results show that the interaction item (MA×DCG) coefficient between managers' ability and digital transformation is 0.031, and it is significantly positive at 1%, which verifies the positive moderating effect of digital transformation (DCG), indicating that digital transformation enhances the promotion effect of managers' ability on the risk-taking level of small and medium-sized enterprises, and hypothesis 4 is verified. Reducing operational risk requires the enterprise to obtain more information, thus reducing the probability of decision deviation. Digital transformation can help enterprises to improve their information processing capabilities, especially the ability to explore and analyze non-standardized data, which can give full play to the regulatory role of enterprise information processing systems in business decisions and production processes. Under the support of digital technology, enterprises promote their decision-making accuracy by absorbing new market information, and improve their risk-taking ability by correcting their deviations.

### 4.4 Robustness Test

In order to make the research conclusion more robust and reliable, the robustness test is carried out by replacing variables and reducing years respectively. Based on the practice of Peter and other scholars[7], the variable replacement method is to sort the regression residuals into four groups from small to large, and measure the manager's ability (MA4) by assigning it to 1, 2, 3 and 4. The higher the assignment, the stronger the manager's ability. Re-substitution into the above model for regression analysis. Shortening the year is to choose the range of 2012-2021, and the regression results all verify the above assumptions again. Tables 4 and 5 are the regression results of robustness test.

**Table 4 Regression Results of Alternative Variables**

<table>
<thead>
<tr>
<th>Variables</th>
<th>(1)Org</th>
<th>(2)RiskT</th>
<th>(3)Org</th>
<th>(4)RiskT</th>
</tr>
</thead>
<tbody>
<tr>
<td>MA4</td>
<td>0.035***</td>
<td>0.002***</td>
<td>0.020***</td>
<td>0.002***</td>
</tr>
<tr>
<td></td>
<td>(18.35)</td>
<td>(32.65)</td>
<td>(6.81)</td>
<td>(12.23)</td>
</tr>
<tr>
<td>RiskT</td>
<td></td>
<td>4.902***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(51.18)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DCG</td>
<td></td>
<td>-0.001***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-2.85)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MA4×DCG</td>
<td></td>
<td>0.027***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3.95)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Controls</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>cons</td>
<td>4.588***</td>
<td>0.047***</td>
<td>4.410***</td>
<td>0.048***</td>
</tr>
<tr>
<td></td>
<td>(116.19)</td>
<td>(36.34)</td>
<td>(83.69)</td>
<td>(34.63)</td>
</tr>
<tr>
<td>N</td>
<td>15801</td>
<td>15801</td>
<td>15801</td>
<td>15801</td>
</tr>
<tr>
<td>Chi-squared</td>
<td>30493.726***</td>
<td>30815.167***</td>
<td>79916.220***</td>
<td>38727.582***</td>
</tr>
</tbody>
</table>
Table 5 Regression Results of Alternative Variables

<table>
<thead>
<tr>
<th>Variables</th>
<th>(1)Org</th>
<th>(2)RiskT</th>
<th>(3)Org</th>
<th>(4)RiskT</th>
</tr>
</thead>
<tbody>
<tr>
<td>MA</td>
<td>0.248***</td>
<td>0.018***</td>
<td>0.103***</td>
<td>0.011***</td>
</tr>
<tr>
<td></td>
<td>(17.74)</td>
<td>(40.11)</td>
<td>(5.35)</td>
<td>(13.76)</td>
</tr>
<tr>
<td>RiskT</td>
<td>4.460***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(48.19)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DCG</td>
<td>0.000***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(4.82)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MA×DCG</td>
<td>0.003***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(9.59)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Controls</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>cons</td>
<td>5.081***</td>
<td>0.039***</td>
<td>4.967***</td>
<td>0.041***</td>
</tr>
<tr>
<td></td>
<td>(115.08)</td>
<td>(55.94)</td>
<td>(102.04)</td>
<td>(56.26)</td>
</tr>
<tr>
<td>N</td>
<td>14323</td>
<td>14323</td>
<td>14323</td>
<td>14323</td>
</tr>
<tr>
<td>Chi-squared</td>
<td>99863.209***</td>
<td>191653.25***</td>
<td>63404.523***</td>
<td>205167.20***</td>
</tr>
</tbody>
</table>

5 CONCLUSIONS AND ENLIGHTENMENT

In the difficult recovery of the global economy, how to improve the ability of managers to improve the level of corporate risk-taking and enhance the organizational resilience of small and medium-sized enterprises has become a very important issue in the research on the sustainable development of small and medium-sized enterprises. This paper uses data envelopment analysis to measure managers' ability, and uses the data of Shanghai and Shenzhen A-share small and medium-sized enterprises from 2010 to 2021 to empirically test the relationship among managers' ability, enterprise risk-taking level and organizational resilience of small and medium-sized enterprises, and the impact of digital transformation on managers' ability and enterprise risk-taking level. The following conclusions and inspirations are drawn: (1) The ability of managers is positively correlated with the firm's resilience level. Improving the ability of managers is beneficial to improving the organizational resilience level of small and medium-sized enterprises. (2) to increase the ability of managers is helpful to improve the risk-taking level of SMEs. (3) The level of risk-taking plays an intermediary role in the influence of managers' ability on the organizational resilience of small and medium-sized enterprises. (4) Digital transformation is significantly adjusting the relationship between managers' ability and risk-taking level of small and medium-sized enterprises, that is, the degree of digital transformation of enterprises can significantly promote the relationship between them. Therefore, in the appointment of managers in small and medium-sized enterprises, attention should be paid to the selection of managers, focusing on the selection of managers with high risk awareness, adaptability and risk-taking level. In addition, it is strengthening daily training and learning, giving more incentives to attract and retain talents, encouraging managers to take risks, enhancing the core competitiveness of enterprises and laying a solid foundation for the sustainable development of small and medium-sized enterprises. In addition, in order to cultivate the crisis awareness of the company team, the formation of the organizational resilience of small and medium-sized enterprises is by no means a last-minute cramming or a brainwave when the crisis comes, or it depends on the super-high resilience of managers, but is the result of long-term accumulation and cultivation, which requires the collective efforts of team members. Finally, clarify the transformation
objectives and gradually promote the pace of digital transformation. At present, small and medium-sized enterprises are relatively lack of digital marketing capabilities and information security measures. Therefore, it is necessary to clarify the transformation objectives and gradually promote the process of digital transformation. Through the advanced digital technology to achieve cost reduction and efficiency, optimize the allocation of resources, social coordination and improve value creation, etc., to ensure the authenticity and symmetry of information, to provide a basic guarantee for the improvement of managers' ability, so as to improve the market competitiveness of small and medium-sized enterprises.

REFERENCES

Research on the Influence Mechanism of R&D Investment and Enterprise Performance Based on FEM Algorithm

Juanjuan Liang1, Linrong Wu1,2,*

*Correspondence should be addressed to Linrong Wu; wulinrong@gxljc.edu.cn

School of Economic and Management, Guilin University, Guilin, 541006, China1
College of Business Administration and Finance Accounting, Rattana Bundit University, Bangkok, 10700, Thailand2

Abstract: In order to reduce the impact of external environment changes on enterprise R&D and performance since 2020, this paper selects the listed companies in Shanghai and Shenzhen A-share market from 2015 to 2019 as the research sample, uses the FEM Algorithm to empirically test the relationship between R&D investment and enterprise performance, and discusses the regulatory role of the heterogeneity of senior management team on it, and adopts the replacement variable measurement method to pass the robustness test. The research finds that R&D investment has a significant negative correlation with the current performance of enterprises and the performance of enterprises lagging behind the first phase, and a significant positive correlation with the performance of enterprises lagging behind the second phase. In the study of the influence of the heterogeneity of senior management team on the relationship between R&D investment intensity and enterprise performance, the heterogeneity of age, career background and education level is positively correlated with the two, which is not conducive to the improvement of enterprise performance. However, the gender and tenure heterogeneity of senior executives and the two have a negative regulatory effect, promoting the improvement of corporate performance.

Keywords: R&D Investment, Enterprise Performance, Heterogeneity of Senior Management Team.

1 INTRODUCTION

Innovation is the source of national sustainable development and a strong strategic support for economic growth. All developed economies in the world are guided by national strategic needs, driven by innovation, relying on high-level scientific and technological innovation, and accumulating strength to carry out original and leading scientific and technological breakthroughs. As an important implementation subject of scientific and technological innovation, enterprises should have long-term core competitiveness and R&D investment is essential. Through R&D investment, enterprises can obtain patents and technologies protected by laws such as patent rights, and their products are favored by consumers and earn excess profits. Secondly, R&D investment can promote enterprises to optimize production mode,
improve efficiency and reduce production costs. Differentiated products also help enterprises to stand out in the market and gain excess market share.

Among the factors affecting the relationship between R&D input and enterprise performance, the key one is the decision-making opinions of the senior management team. Research and development investment is characterized by long investment time, high risk and uncertain returns. From the perspective of agency theory, senior executives may pay little attention to research and development investment with uncertain returns based on short-term benefits. The level of decision-making is also influenced by the differing opinions of the senior management team.

The main contributions of this paper are as follows: First, through empirical testing of the impact of R&D input on the corporate performance of Shanghai and Shenzhen A-share listed companies during 2015-2019 and the following two periods, the research gap of the relationship between R&D input and corporate performance in recent years was filled; Second, the five dimensions of the heterogeneity of the top management team are studied to measure the impact of the heterogeneity of the top management team in different dimensions on the relationship between R&D input and corporate performance, which has practical significance for the innovation implementation of listed companies.

2 THEORETICAL ANALYSIS AND RESEARCH HYPOTHESIS

2.1 R&D Investment and Enterprise Performance

The relationship between R&D investment and enterprise performance in the current period. Ben Branch (1974) analyzed 111 companies in the United States from 1950 to 1965 and found that, except for the pharmaceutical industry, R&D investment in other six industries promoted the growth of corporate profitability. Hitt, Hoskisson (1991) studied 191 companies in the United States from 1970 to 1986 and found that R&D investment negatively regulated enterprise performance among companies across 29 industries. Lantz J S, Sahut J M (2005) found through the empirical model that if the R&D investment can obtain income in the current period, its capitalization can increase financial performance, if it can obtain income in the future, the reverse is true. The relationship between R&D investment and enterprise performance is nonlinear. Alam A, et al (2020) used GMM method to study 432 companies in 12 countries, and found that safeguard measures effectively regulate the positive correlation between R&D investment and enterprise performance. Xu J, Wang X, Liu F (2021) believed that R&D investment has a positive and significant effect on performance, especially for state-owned enterprises and enterprises with R&D directors. The R&D investment will make the enterprise ahead of its peers in terms of products, production costs and market share.

Based on this, this paper proposes:

Hypothesis 1. R&D investment will promote the current performance of enterprises.

The relationship between R&D investment and enterprise performance lags behind. William (1971) took the U.S. manufacturing industry from 1957 to 1967 as the research sample, and found that R&D investment has a positive lag effect on enterprise performance, which is
reflected in the continuous growth of sales revenue and profit in the second to ninth years. Chambers (2002) found that the lagged relationship between R&D investment and enterprise performance has lasted for about ten years. This conclusion was drawn through empirical research on nearly 10,000 American enterprises. Zhu Z, Huang F (2012) conducted research on listed technology-based enterprises and found that there is a positive correlation between R&D investment and enterprise performance in the next year. Li Lu and Zhang Wanting (2013) believed that R&D investment positively adjusted the current performance, and there was a lag effect of two periods, which decreased year by year. The results of R&D investment and market audience feedback are gradually realized, but how to achieve the effect remains to be verified.

Based on this, this paper proposes:

**Hypothesis 2.** R&D investment will promote the current performance of enterprises.

### 2.2 The Influence of the Heterogeneity of Senior Management Team on the Correlation Between R&D Investment and Enterprise Performance

As decision-makers of major business strategies, senior executives play a key role in R&D investment and performance of enterprises. In the background of agency theory, different senior management team characteristics play a differentiated role in the final business decisions and the realization of corporate goals.

Regulation of age heterogeneity. Zenger and Lawrence (1989) believed that the greater the age heterogeneity, the less technical communication within the project team. Richard O C, Shelor R M (2002) found that the middle and low level of age heterogeneity promoted sales growth, while the high level of age heterogeneity inhibited sales growth. Wang H, He W, Yang Y (2022) said that in innovation-oriented enterprises, age heterogeneity will not affect enterprise innovation performance, and more attention should be paid to team members' abilities, personality, and work adaptability. The senior management team with significant age difference is more conducive to coping with the complex internal and external environment.

Based on this, this paper proposes:

**Hypothesis 3.** The age heterogeneity of senior executives has a positive moderating effect on R&D investment and enterprise performance

Regulation of gender heterogeneity. Rivero, Arlene (2003) believed that female executives, compared with male executives, had keen insight and communication skills, and would make efficient decisions more quickly using existing information and resources in the face of risk decisions. Dezso (2011) empirically found that the participation of senior female executives was positively correlated with corporate performance. Ren Ting et al. (2010) believed that the increase in the proportion of female executives in the team is positively improving enterprise performance. In terms of innovation decision-making, female executives are cautious. Anderson R C, Reeb D M (2011) found that in the board of directors, gender differences help improve the company's performance and bring in different talents and perspectives.

Based on this, this paper proposes:

**Hypothesis 4.** The gender heterogeneity of senior executives has a positive moderating effect on R&D investment and enterprise performance.
The moderating effect of occupational background heterogeneity. Sutcliffe (1994) believes that the diversity of professional backgrounds of senior management team members will reduce the efficiency of decision-making communication, weaken the ability to identify external opportunities, and be unfavorable to making effective decisions to improve corporate performance. Simons (1999) believed that senior management teams with different career backgrounds can enrich the team's cognition, skills and handling methods, which is conducive to the making of corporate performance decisions. Ndofor H A, Simron D G, He X (2015) believed that occupational background heterogeneity promoted the resource-action connection, but had a negative impact on the action-performance connection.

Based on this, this paper proposes:

*Hypothesis 5.* The heterogeneity of senior executives' occupational background has a negative moderating effect on R&D investment and firm performance.

The moderating effect of educational level heterogeneity. Smith (1994) believed that the more differentiated the education level of the top management team, the deeper understanding of the phenomenon, is conducive to improving the quality of decision-making and corporate performance. Pinelli M, Cappa F, Franco S, et al (2020) studied the education level and educational background of the founders of 1078 start-ups and found that the heterogeneity of education level and educational background promoted the increase of the amount of enterprise financing, but the coexistence of both weakened the positive relationship. In-depth information mining and cognition requires the support of high education level.

Based on this, this paper proposes:

*Hypothesis 6.* The heterogeneity of executive education level has a positive moderating effect on R&D investment and firm performance.

The moderating role of tenure heterogeneity. Dutton (1987) believed that senior management teams with different terms of office have a comprehensive perspective on information collection and interpretation, which is more conducive for enterprises to make reasonable and optimized decisions. Boeker (1997) believed that senior management teams with large tenure differences have more opportunities to adopt different management styles and innovative strategic plans. Wang H, He W, Yang Y (2022) should try to maintain the same tenure of senior management team and extend the tenure, which is conducive to communication and exchange, and improve the efficiency and efficiency of decision-making. The length of managers' tenure affects their familiarity with the operation of enterprise management mode and employees.

Based on this, this paper proposes:

*Hypothesis 7.* The heterogeneity of executive tenure has a negative moderating effect on R&D investment and firm performance.
3 RESEARCH DESIGN

3.1 Sample Selection and Data Source

According to the existing research, the research samples in this paper are from the relevant data of Shanghai and Shenzhen A-share listed companies from 2015 to 2019, excluding PT, ST, *ST companies, financial companies and companies with abnormal data. The sample data is mainly from the CSMAR database, and the missing value is manually collected from the enterprise annual report. Data processing was mainly completed by stata16, and was processed by 1% horizontal indentation.

3.2 Variable Definition

Explained variable: firm performance. Based on the existing research literature, as well as the availability and quantification of data, return on assets (ROA) is used as a variable to measure corporate performance. Return on Assets (ROA) is the ratio of a company's net profit to the average amount of assets, which can be obtained by the average of the ending balance of assets.

Explanatory variable: R&D input Rd. In the past, a quantifiable measure of R&D investment was divided by total assets, operating profit, net profit, etc. In this paper, the ratio of R&D investment to operating revenue is adopted to measure the degree of R&D investment, which is studied by Zhang Zhaoguo et al. (2014). In order to study the impact of R&D input on enterprise benefits in the current period and lag period, the data of R&D input in the current period, lag phase I and lag phase II are taken in this paper.

Moderating variable: Heterogeneity of top management team. This paper measures the heterogeneity of senior management team from five dimensions: age Hage, sex Hsex, education level Hdegree, career background Hbackground and tenure Hterm. Age Hage and tenure heterogeneity Hterm are continuous variables, measured by the ratio of standard deviation to mean. Gender Hsex, education level Hdegree and career background Hbackground were measured by Hersman coefficient $H = 1 - \sum_i P_i^2$, where $P_i$ is the percentage of Class $i$ members in the team, $n$ is the number of types, and the larger the value of $H$, the higher the degree of heterogeneity.

Control variables: referring to Li Xianjun (2018) et al. and Wang Xi (2020) et al., this paper introduces the control variables of enterprise growth, asset liability ratio, enterprise size Lnassets and enterprise age Age. See Table 1 for variable settings.

<table>
<thead>
<tr>
<th>Table1 Variable Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type</strong></td>
</tr>
<tr>
<td>Explained variable</td>
</tr>
<tr>
<td>Explanatory variable</td>
</tr>
<tr>
<td>Adjustment variable</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
### 3.3 Model Setting

(1) The impact of R&D investment on firm performance. In order to study the impact of R&D input on enterprise performance in the current period and its lag time, three regression models of R&D input on enterprise performance in the current period, one lag time and two lag time were established.

\[
\text{Roai}_i = r + \alpha_1 \text{Rd}_{i,t} + \alpha_2 \text{Growth}_{i,t} + \alpha_3 \text{Ratio}_{i,t} + \alpha_4 \ln \text{assets}_{i,t} + \alpha_5 \text{Age}_{i,t} + \epsilon_{i,t} \quad (1)
\]

\[
\text{Roai}_i = r + \alpha_1 \text{Rd}_{i,t-1} + \alpha_2 \text{Growth}_{i,t} + \alpha_3 \text{Ratio}_{i,t} + \alpha_4 \ln \text{assets}_{i,t} + \alpha_5 \text{Age}_{i,t} + \epsilon_{i,t} \quad (2)
\]

\[
\text{Roai}_i = r + \alpha_1 \text{Rd}_{i,t-2} + \alpha_2 \text{Growth}_{i,t} + \alpha_3 \text{Ratio}_{i,t} + \alpha_4 \ln \text{assets}_{i,t} + \alpha_5 \text{Age}_{i,t} + \epsilon_{i,t} \quad (3)
\]

(2) The moderating influence of the heterogeneity of top management team on R&D investment and firm performance. On the basis of model (1), the interaction term between the heterogeneity of the top executive team and R&D investment is added to determine whether the heterogeneity of the top executive team has a moderating effect on the corporate performance of R&D investment and the degree of influence.

\[
\text{Roai}_i = r + \alpha_1 \text{Hage} \cdot \text{Rd}_{i,t} + \alpha_2 \text{Rd}_{i,t} + \alpha_3 \text{Growth}_{i,t} + \alpha_4 \text{Ratio}_{i,t} + \alpha_5 \ln \text{assets}_{i,t} + \alpha_6 \text{Age}_{i,t} + \epsilon_{i,t} \quad (4)
\]

\[
\text{Roai}_i = r + \alpha_1 \text{Hsex} \cdot \text{Rd}_{i,t} + \alpha_2 \text{Rd}_{i,t} + \alpha_3 \text{Growth}_{i,t} + \alpha_4 \text{Ratio}_{i,t} + \alpha_5 \ln \text{assets}_{i,t} + \alpha_6 \text{Age}_{i,t} + \epsilon_{i,t} \quad (5)
\]

\[
\text{Roai}_i = r + \alpha_1 \text{Hbackground} \cdot \text{Rd}_{i,t} + \alpha_2 \text{Rd}_{i,t} + \alpha_3 \text{Growth}_{i,t} + \alpha_4 \text{Ratio}_{i,t} + \alpha_5 \ln \text{assets}_{i,t} + \alpha_6 \text{Age}_{i,t} + \epsilon_{i,t} \quad (6)
\]

\[
\text{Roai}_i = r + \alpha_1 \text{Hdegree} \cdot \text{Rd}_{i,t} + \alpha_2 \text{Rd}_{i,t} + \alpha_3 \text{Growth}_{i,t} + \alpha_4 \text{Ratio}_{i,t} + \alpha_5 \ln \text{assets}_{i,t} + \alpha_6 \text{Age}_{i,t} + \epsilon_{i,t} \quad (7)
\]

\[
\text{Roai}_i = r + \alpha_1 \text{Hterm} \cdot \text{Rd}_{i,t} + \alpha_2 \text{Rd}_{i,t} + \alpha_3 \text{Growth}_{i,t} + \alpha_4 \text{Ratio}_{i,t} + \alpha_5 \ln \text{assets}_{i,t} + \alpha_6 \text{Age}_{i,t} + \epsilon_{i,t} \quad (8)
\]
4 RESULTS AND DISCUSSION

4.1 Descriptive Statistics and Correlation Analysis

The descriptive statistical characteristics of variables in this paper are shown in the Table 2. From the descriptive statistics of 13172 observations, we can see that: (1) The corporate performance of listed companies varies greatly, the maximum is 0.199, the minimum is -0.351, and the average is 0.038. Most of the corporate performance is positive. (2) The average R&D investment intensity of A-share listed companies in the past five years has reached 4.719%, which is close to the internationally competitive R&D investment standard of 5%, indicating that Chinese listed companies attach importance to R&D to build their own core competitiveness, and building innovation brings commercial moats. (3) The average of the heterogeneity of education level and occupational background of the senior management team is 0.23 and 0.21, respectively, indicating that the education level and occupational background of the senior management team are quite different. Besides, this study conducts Pearson test on each variable. From Table 3, it can be seen that the correlation coefficient between R&D investment and enterprise performance variable is -0.038, which is negative significant at the level of 1%. This means that the increase of R&D investment will lead to the decrease of enterprise performance in the current period. The assumption H1 is preliminarily rejected. Next, we made a collinearity diagnosis for all variables. From Table 4, we can see that the variance expansion factor is less than 10, and there is no high correlation between the variables. On this basis, Hausman test was conducted, and each model rejected the original hypothesis, and the regression analysis method selected the FEM Algorithm.

Table 2 Descriptive statistics of the variables

<table>
<thead>
<tr>
<th>variable</th>
<th>N</th>
<th>mean</th>
<th>p50</th>
<th>sd</th>
<th>min</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roa</td>
<td>13172</td>
<td>0.0380</td>
<td>0.0410</td>
<td>0.0710</td>
<td>-0.351</td>
<td>0.199</td>
</tr>
<tr>
<td>Rd</td>
<td>13172</td>
<td>4.719</td>
<td>3.690</td>
<td>4.496</td>
<td>0.0300</td>
<td>25.37</td>
</tr>
<tr>
<td>Hage</td>
<td>13172</td>
<td>0.129</td>
<td>0.124</td>
<td>0.0520</td>
<td>0.0300</td>
<td>0.277</td>
</tr>
<tr>
<td>Hsex</td>
<td>13172</td>
<td>0.220</td>
<td>0.200</td>
<td>0.136</td>
<td>0.0560</td>
<td>1</td>
</tr>
<tr>
<td>Hbackground</td>
<td>13172</td>
<td>0.210</td>
<td>0.195</td>
<td>0.0940</td>
<td>0.0650</td>
<td>0.595</td>
</tr>
<tr>
<td>Hdegree</td>
<td>13172</td>
<td>0.230</td>
<td>0.176</td>
<td>0.208</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Hterm</td>
<td>13172</td>
<td>0.595</td>
<td>0.590</td>
<td>0.316</td>
<td>0</td>
<td>1.430</td>
</tr>
<tr>
<td>Growth</td>
<td>13172</td>
<td>0.178</td>
<td>0.116</td>
<td>0.377</td>
<td>-0.511</td>
<td>2.330</td>
</tr>
<tr>
<td>Ratio</td>
<td>13172</td>
<td>0.400</td>
<td>0.385</td>
<td>0.197</td>
<td>0.0600</td>
<td>0.912</td>
</tr>
<tr>
<td>Inassets</td>
<td>13172</td>
<td>22.15</td>
<td>21.98</td>
<td>1.268</td>
<td>19.96</td>
<td>26.18</td>
</tr>
<tr>
<td>Age</td>
<td>13172</td>
<td>8.926</td>
<td>7</td>
<td>7.341</td>
<td>-1</td>
<td>25</td>
</tr>
</tbody>
</table>

Table 3 Correlation Matrix

<table>
<thead>
<tr>
<th>variable</th>
<th>Roa</th>
<th>Rd</th>
<th>Hage</th>
<th>Hsex</th>
<th>Hdegree</th>
<th>Hterm</th>
<th>Growth</th>
<th>Ratio</th>
<th>Inassets</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roa</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rd</td>
<td>-</td>
<td>0.038*</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hage</td>
<td>-0.013</td>
<td>0.049*</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Regression Analysis

(1) Influence of R&D investment on enterprise performance

From the regression results in Table 5, the regression coefficient of R&D investment on the current and lagging corporate performance is -0.001, which is significant at the level of 1%, indicating that R&D investment has hindered the improvement of corporate financial performance in the current and lagging corporate performance, so we refuse to assume Hypothesis 1. However, the enterprise performance coefficient of R&D investment and lag phase II is 0.003, which is significant at the level of 1%, which means that R&D input-output
efficiency is positively promoting the enterprise performance of lag phase II. Based on the regression results of lag phase I and lag phase II, the relationship between R&D investment and enterprise performance is lagging behind. Hypothesis 2 is verified. Analyze the reasons. The benefits brought by the R&D investment in the current period and the first phase lag behind failed to make up for the costs paid. The second phase lag behind turns into positive benefits, indicating that the R&D investment can bring positive benefits. This is different from most research results such as Liang Laixin, Zhang Huanfeng (2005), Ren Haiyun (2009), and Zhang Jian (2014). It may be due to the differences in the years of sample data research, as well as different research industries and sectors. Combining the descriptive statistics above, it also verifies that China's listed companies pay more attention to R&D investment, increasing from the international pass line of 2% to 4.821%.

Table 5 Regression Results of the Current and Lagging Effects of R&D Investment on Enterprise Performance

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model (1)</th>
<th>Model (4)</th>
<th>Model (5)</th>
<th>Model (6)</th>
<th>Model (7)</th>
<th>Model (8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rd</td>
<td>-0.001***</td>
<td>-0.007***</td>
<td>-0.005***</td>
<td>-0.006***</td>
<td>-0.005***</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Growth</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Ratio</td>
<td>-0.351***</td>
<td>-0.346***</td>
<td>-0.351***</td>
<td>-0.350***</td>
<td>-0.350***</td>
<td>-0.350***</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.008)</td>
<td>(0.008)</td>
<td>(0.008)</td>
<td>(0.008)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>Lnassets</td>
<td>0.061***</td>
<td>0.057***</td>
<td>0.061***</td>
<td>0.060***</td>
<td>0.060***</td>
<td>0.061***</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Age</td>
<td>-0.011***</td>
<td>-0.010***</td>
<td>-0.010***</td>
<td>-0.010***</td>
<td>-0.010***</td>
<td>-0.010***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Hage*Rd</td>
<td>0.019***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hsex*Rd</td>
<td></td>
<td>-0.039***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hbackground*Rd</td>
<td></td>
<td></td>
<td>0.007***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.001)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hdegree*Rd</td>
<td></td>
<td></td>
<td></td>
<td>0.005***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hterm*Rd</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.005***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td>_cons</td>
<td>-1.072***</td>
<td>-0.960***</td>
<td>-1.048***</td>
<td>-1.031***</td>
<td>-1.036***</td>
<td>-1.070***</td>
</tr>
<tr>
<td></td>
<td>(0.085)</td>
<td>(0.085)</td>
<td>(0.085)</td>
<td>(0.085)</td>
<td>(0.085)</td>
<td>(0.085)</td>
</tr>
<tr>
<td>N</td>
<td>13191.000</td>
<td>13172.000</td>
<td>13186.000</td>
<td>13186.000</td>
<td>13186.000</td>
<td>13188.000</td>
</tr>
</tbody>
</table>
(2) The moderating effect of TMT heterogeneity on R&D investment on firm performance

In order to study the impact of the heterogeneity of the senior management team on R&D investment and enterprise performance, this paper sets cross terms Hage * Rd, Hsex * Rd, Hbackground * Rd, Hdegree * Rd and Hterm * Rd to test the moderating effects of the heterogeneity of senior management age, gender, professional background, education level and tenure. The regression results are shown in Table 6. All interaction items are significant, which means that the heterogeneity of the senior management team has a moderating effect. The Hage * Rd coefficient of the interaction term of senior executives' age heterogeneity is 0.019, which is significant at the level of 1%, indicating that senior executives' age heterogeneity will strengthen (weaken) the negative (positive) relationship between R&D investment and enterprise performance, and there is a positive moderating effect. H3 is verified. It can be seen from the regression coefficient that the interaction cross term Hsex * Rd of senior executives' gender heterogeneity negatively regulates the relationship between the two, assuming that H4 refuses to verify. The heterogeneous interaction term Hbackground * Rd of senior executives' professional background and the heterogeneous interaction term Hdegree * Rd of senior executives' education level are positive significant at the level of 1%, indicating that these two will strengthen (weaken) the negative (positive) relationship between R&D investment and enterprise performance, and have a positive regulatory effect. Therefore, we refuse to assume H5, and assume H6 is verified. The heterogeneity interaction term Hterm * Rd of senior executives' tenure is negative significant, which means that the greater the diversity of senior executives' tenure, the less the effect of R&D investment on corporate performance will be. Hterm * Rd has negative regulation effect, and H7 is assumed to be verified.

### Table 6 The moderating effect of the heterogeneity of senior management team on the correlation between R&D investment and enterprise performance

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model (1)</th>
<th>Model (4)</th>
<th>Model (5)</th>
<th>Model (6)</th>
<th>Model (7)</th>
<th>Model (8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rd</td>
<td>-0.001***</td>
<td>-0.007***</td>
<td>-0.005***</td>
<td>-0.006***</td>
<td>-0.005***</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Growth</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Ratio</td>
<td>-0.351***</td>
<td>-0.346***</td>
<td>-0.351***</td>
<td>-0.350***</td>
<td>-0.350***</td>
<td>-0.350***</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.008)</td>
<td>(0.008)</td>
<td>(0.008)</td>
<td>(0.008)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>Lnassets</td>
<td>0.061***</td>
<td>0.057***</td>
<td>0.061***</td>
<td>0.060***</td>
<td>0.060***</td>
<td>0.061***</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Age</td>
<td>-0.011***</td>
<td>-0.010***</td>
<td>-0.010***</td>
<td>-0.010***</td>
<td>-0.010***</td>
<td>-0.010***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Hage*Rd</td>
<td>0.019***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hsex*Rd</td>
<td></td>
<td>-0.039***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.013)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hbackground*Rd</td>
<td></td>
<td></td>
<td>0.007***</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: *, ** and *** respectively represent 10%, 5% and 1% significance, and t value is in brackets.
4.3 Robustness Test

In order to avoid the contingency of the above regression results, this paper uses Eps earnings per share to replace the enterprise performance Roa variable to verify the impact of R&D investment on enterprise performance, and the moderating effect of the heterogeneity of senior management team on both. It can be seen from Table 7 that after replacing earnings per share to measure enterprise performance, the conclusion that R&D investment lags behind one and two periods of enterprise performance in the current period is still valid. Therefore, the negative significant impact of R&D investment on enterprise performance is still valid, and the regulatory effects of the heterogeneity of senior management team are consistent, indicating that the study passed the robustness test.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model (1)</th>
<th>Model (4)</th>
<th>Model (5)</th>
<th>Model (6)</th>
<th>Model (7)</th>
<th>Model (8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rd</td>
<td>-0.007***</td>
<td>-0.036***</td>
<td>-0.026***</td>
<td>-0.032***</td>
<td>-0.024***</td>
<td>-0.002</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.005)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Growth</td>
<td>0.001**</td>
<td>0.001**</td>
<td>0.001**</td>
<td>0.001**</td>
<td>0.001**</td>
<td>0.001**</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Ratio</td>
<td>-0.917***</td>
<td>-0.932***</td>
<td>-0.912***</td>
<td>-0.909***</td>
<td>-0.909***</td>
<td>-0.914***</td>
</tr>
<tr>
<td></td>
<td>(0.040)</td>
<td>(0.040)</td>
<td>(0.040)</td>
<td>(0.040)</td>
<td>(0.040)</td>
<td>(0.040)</td>
</tr>
<tr>
<td>Lnassets</td>
<td>0.371***</td>
<td>0.368***</td>
<td>0.369***</td>
<td>0.367***</td>
<td>0.366***</td>
<td>0.368***</td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td>(0.021)</td>
<td>(0.020)</td>
<td>(0.020)</td>
<td>(0.020)</td>
<td>(0.020)</td>
</tr>
<tr>
<td>Age</td>
<td>-0.047***</td>
<td>-0.045***</td>
<td>-0.044***</td>
<td>-0.044***</td>
<td>-0.045***</td>
<td>-0.044***</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>Hage*Rd</td>
<td>0.096***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.030)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hsex*Rd</td>
<td></td>
<td>0.027***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.004)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hbackground<em>rd</em>Rd</td>
<td></td>
<td>0.035***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.004)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hdegree*Rd</td>
<td></td>
<td></td>
<td>0.024***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.004)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hterm*Rd</td>
<td></td>
<td></td>
<td></td>
<td>-0.020***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.005)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
5 CONCLUSIONS

First of all, the R&D investment of enterprises is significantly negatively correlated with the enterprise performance of the current period and the first lag period, and significantly positively correlated with the enterprise performance of the second lag period. R&D investment weakens the enterprise performance in the current period and the following year, but it will significantly improve the enterprise performance in the following two years, which is different from the conclusions of most scholars. The reasons are: first, the research expenditure in the early stage of R&D is included in the current profit and loss. Although there are tax incentives for R&D, it still has a negative impact on the current performance [29]; second, the technology acquired by R&D investment and the intangible assets transformed need time to pave the way, which has a negative impact on the performance of enterprises lagging behind Phase I, but ultimately has a positive effect on the performance of enterprises lagging behind Phase II.

Secondly, the heterogeneity of senior management team has a moderating effect on the relationship between R&D investment intensity and enterprise performance. The heterogeneity of senior executives' age, professional background and education level is positively related to the two, which is not conducive to the improvement of enterprise performance. The gender and tenure heterogeneity of senior executives and the two are negative moderating effects, promoting enterprise performance. The enterprise can adjust the characteristics of the senior management team in a timely manner according to the current relationship between R&D investment and enterprise performance, in order to optimize enterprise performance.

Funding Statement: This research is supported by the basic research promotion project for young and middle-aged teachers in Guangxi's colleges and universities, "Research and Development Investment, Heterogeneity of Senior Management Teams and the Impact Mechanism of Enterprise Performance and Empirical Research" (2022KY1570)

REFERENCES


Research on the Transformation of Financial Accounting to Management Accounting under the Background of Big Data Combined with Algorithm Model

Chao Li
624299703@qq.com
School of accounting, Wuhan Qingchuan University, Hongshan District, Wuhan City, Hubei Province, China

Abstract: With the advent of the era of big data, big data provides technical support for enterprises' financial activities, and produces massive data, which has a great influence on our country's economic development. The transformation of financial accounting to management accounting based on big data technology has become the development trend of the information age and is also the inevitable requirement of enterprise development. Based on this, this paper firstly collects and processes the required data using the statistical algorithm model, and then analyses the significance of the transformation from financial accounting to management accounting under the background of big data. Finally, the paper concludes that 1) enterprises need to guide financial personnel, completely change the traditional work concept, to ensure that management accounting work can meet the needs of modern development of enterprises. 2) Only the integration of financial accounting and management accounting can reduce the consumption of financial and human resources and improve the internal management system. 3) On the basis of the introduction of advanced information platform, innovate management accounting concepts, ensure the rationality and authenticity of data, and significantly improve the economic value of enterprises.

Keywords: Big Data, Financial Accounting, Management Accounting, Algorithm Model, Risk Control.

1 INTRODUCTION

With the continuous development of information technology, enterprises need to shift from financial accounting to management accounting in order to get further development and play the application significance of big data [3, 4, 8]. As an important part of the enterprise, the financial management department is affected by the development of information technology [5, 12, 15]. The work and management form of financial accounting can no longer meet the needs of enterprise management. In order to change this situation, it is necessary for managers to clarify the development direction of the big data background and accelerate the transformation of management accounting [2, 10, 13]. However, in the process of transformation and adjustment, managers need to actively guide and help accounting personnel learn information technology, optimize the working mode of management accounting, achieve the purpose of improving the
efficiency of information resource integration, and promote the operation and development of enterprises [5, 10, 13].

The era of big data has the characteristics of large amount of data, strong authenticity, and can carry out centralized data processing in a short time [1, 11]. In the actual operation process, enterprises will produce a large amount of data. Only reasonable use of data processing technology to ensure the integrity of data information can provide a basis for various business decisions [1, 7, 9]. Enterprises realize the transformation and development of financial accounting to management accounting, timely find the shortcomings of financial control, pay attention to the whole process of risk control, strengthen the capital supervision and management, can play the basic function of financial control, avoid unnecessary financial risks and business risks, improve the core competitiveness of enterprises [3, 7, 9].

This paper analyses the significance of the transformation from financial accounting to management accounting under the background of big data, and puts forward several measures to innovate the concept of accounting management data, realize the integration of financial accounting and management accounting, and establish a big data financial information platform, hoping to provide reference for relevant people.

2 ALGORITHM MODEL [14]

In order to better mine the data, this paper uses K harmonic mean arithmetic (KHM) clustering algorithm to collect the data.

2.1 Introduction for Algorithm Model

The following is a brief introduction to the basic information of the algorithm:

KHM clustering algorithm is based on central iterative process. Assuming that \( Z = \{Z_1, Z_2, ..., Z_n\} \) is a set of data, including \( Z_i = \{Z_{i,1}, Z_{i,2}, ..., Z_{i,m}\} \) said with \( m \) an attribute of a data object, cluster number \( k \), the clustering centre set to \( C = \{C_1, C_2, ..., C_k\} \). The following are the main differences between KHM clustering algorithm and K-means algorithm:

(1) In KHM algorithm, the minimum distance function in K-means algorithm is replaced by harmonic mean function, and the harmonic mean formula is:

\[
\frac{1}{\sum_{c \in C} \frac{k}{d^2(z, c)}}
\]

where, \( z \in Z \) represents the object in the data set, \( c \in C \) represents the clustering centre, and \( d^2(z, c) \) is the distance measure function.

(2) The objective function of I deletion algorithm is:

\[
E_{KHM} = \sum_{z \in Z} \frac{k}{\sum_{c \in C} \frac{1}{d^2(z, c)}}
\]

where, \( E_{KHM} \) represents the sum of the harmonic mean of all data points to the centre of each cluster.
(3) The central iteration formula is mainly used to update the clustering centre, and its expression is as follows:

\[ c_k = \frac{\sum_{z \in Z} \frac{1}{(\sum_{y \in Y} \frac{d^2(z, c_k)}{d^2(z, y)})^2}}{\sum_{z \in Z} \frac{1}{(\sum_{y \in Y} \frac{d^2(z, c_k)}{d^2(z, y)})^2}} \]

### 2.2 Clustering Result Evaluation Method

F-measure was obtained by synthesizing Precision and Recall in the field of information retrieval. In this paper, F-measure was used to evaluate the clustering results. For known classification \( i \) and arbitrary cluster \( j \), the accuracy and recall ratio are defined as:

\[ P(i, j) = \frac{M_{ij}}{M_i} \]
\[ R(i, j) = \frac{M_{ij}}{M_j} \]

The \( M_i \) in the formula represents the number of objects in class \( i \). \( M_j \) represents the number of objects in cluster \( j \). \( M_{ij} \) represents the number of objects belonging to a given class \( i \) in cluster \( j \). F-measure of classification \( i \) can be defined as:

\[ F(i) = \frac{2PR}{P + R} \]

\( P = P(i, j) \) and \( R = R(i, j) \) in the formula. Think of \( F(i) \) as the system's score for classification \( i \). Consider the cluster with the highest \( F(i) \) value as the corresponding classification \( i \). The total F-measure is the weighted average of each class \( i \):

\[ F - \text{measure} = \frac{\sum_{i \in N} |i| \times F(i)}{\sum_{i \in N} |i|} \]

where, \( N \) represents all known classes, and \( |i| \) represents the number of objects in class \( i \). The larger the F-measure is, the better the clustering effect is.

### 3 THE SIGNIFICANCE OF THE TRANSITION FROM FINANCIAL ACCOUNTING TO MANAGEMENT ACCOUNTING UNDER THE BACKGROUND OF BIG DATA

#### 3.1 Correct Handling of Accounting Information Data

Based on the background of big data, the growth rate of information is accelerating. However, due to the limited financial accounting processing capacity of the enterprise, it is unable to organize the data in a short time, and it cannot meet the requirements of the development of informatization for the enterprise. The introduction of the concept of management accounting can not only deal with financial information, but also deal with non-financial information and
information, and carry out the classification of various information data, so as to formulate a perfect economic development plan. Only by making a reasonable development plan can we improve the level of funds used by enterprises. Relevant staff should be combined with the use of data, plan clear standards, and constantly improve management efficiency. According to the needs of enterprise accounting work is not difficult to know, in order to achieve the restraint of accounting personnel, it is necessary to analyse the qualified standards of accounting positions and ensure the unified management of data content.

3.2 Strengthen Financial Supervision

The integration of information technology and management accounting can improve the accounting management efficiency of enterprises to a certain extent, and constantly strengthen supervision and control. The adoption of cloud accounting services can make financial personnel not restricted by place and time, but strengthen the communication and exchange of various departments, and realize the transformation of financial accounting. While implementing the management accounting concept, I can prepare medium- and long-term development goals. In addition, big data technology is taken as the core content to prepare financial planning programs, thus providing important basis for various production and business activities and conducive to the analysis of financial reports.

3.3 Promote Enterprise Development

Enterprise financial accounting in the application of traditional methods of work, its focus is to statistical financial applications to draw a perfect statement. Managers can understand the status quo of economic development of enterprises through this content, and then adjust operations according to the development of enterprises. Although financial accounting has the function of summarizing data, management accounting can analyse financial information from multiple angles to improve the purpose of fine processing. For example, when sorting out information, management accounting will focus on the recent financial situation and the future development trend of enterprises [5]. By applying this method to financial management, the problems existing in the development of enterprises can be found in time. After understanding this situation, managers can make timely adjustments to avoid financial damage, promote enterprise development and improve the overall level of the enterprise.

3.4 Meet the Development Needs of Financial Accounting

Influenced by the era of big data, all walks of life gradually adjust their work forms. Enterprise financial internal data processing has also taken place a certain change. In the application of traditional methods to process accounting data information, due to the large amount of various information data, information screening and sorting work is more difficult, once the problem of data errors, will directly affect the operation and development of enterprises. However, after integrating big data technology into enterprises and transforming to management accounting, the system can quickly read data information and improve the efficiency of information sorting and classification. This work form has high timeliness and integrity, and can meet the needs of accounting development.
4 THE DIFFERENCE BETWEEN FINANCIAL ACCOUNTING AND MANAGEMENT ACCOUNTING

4.1 Different Work Priorities

In order to further understand the differences between financial accounting and management accounting, it can be discussed from the perspective of different work priorities. Financial accounting needs to take statements as the basis when working. By showing the operation of the enterprise to managers in this way, the focus of financial accounting is to sort out the financial situation of the enterprise. However, the goal of management accounting is to serve the internal management of enterprises, so it is necessary to study the operation and management of enterprises and provide theoretical reference for managers to make decisions. Under the background of big data, financial accounting needs to apply advanced technology to sort out enterprise information data, improve the timeliness of work and the accuracy of data information, and have a clear understanding of enterprise expenditure and income. For instance, we can check 5G direct/indirect economic output scale in 2020-2030 in the Figure 1.

![Figure 1](image)

Figure 1 5G direct/indirect economic output scale in 2020-2030 (Unit: trillion Yuan).

4.2 Different Work Characteristics

The job of financial accounting is to convey financial information. Financial personnel often make all kinds of enterprise financial statements in the work, is also the main work characteristics of financial accounting. And management accounting work, because this work needs to meet all kinds of enterprise management needs. Due to the different formats of information, staff are required to flexibly sort out various types of reports, which makes the work of management accounting characterized by the characteristics of diversified report forms. We could see from Figure 2 that Change Trend of China Satellite Communication Market Size from 2015 to 2020.
THE EFFECTIVE PATH OF TRANSFORMATION FROM FINANCIAL ACCOUNTING TO MANAGEMENT ACCOUNTING UNDER THE BACKGROUND OF BIG DATA

5.1 Innovation of Accounting Management Data Concept, Accelerate the Pace of Accounting Information Construction

Under the background of continuous progress and development of Chinese society, the traditional way of financial accounting is outdated. Due to the changing management requirements of enterprises, the requirements of financial management are not only the preparation of perfect financial statements, records of daily expenditure work, will also require the implementation of budget management and performance management. Because of the direct connection between the functions of financial accounting and management accounting, we can carry out the assessment of various financial work. But the basic function of financial accounting is to carry out accounting work. Management accounting is the content of money as the core, using the way of recording, calculation, measurement, to achieve the record of the consumption of funds in various business activities of enterprises. Therefore, the enterprise management personnel need to analyse the financial accounting management mechanism, combined with financial accounting statements and fund flow situation, intuitive understanding of the financial development of the enterprise. By analysing the change direction of management accounting, managers can make reasonable use of the information platform and implement the integrated management of financial information data, which provides an important basis for the reasonable allocation of funds. By constantly increasing the research on construction projects, various departments can communicate with each other through the information system, effectively accelerate the speed of information flow and transmission, and ensure the rationality of enterprise management decisions.

5.2 Realize the Integration of Financial Accounting and Management Accounting to Implement the Whole Process of Risk Control

Based on the background of the era of big data, the speed and effectiveness of information transmission have put forward clear requirements for the processing of accounting information. Traditional accounting work of enterprises cannot meet the needs of current
strategic development. Only by shortening the time consumed by accountants in the accounting process can we create higher social and economic benefits for enterprises. The work of management accounting pays more attention to the work of forecasting and financial planning, and guarantees the authenticity of accounting data, which is reflected in the following aspects: traditional financial accounting is to prepare financial statements, to summarize and process the financial data of each year, ignoring the analysis of the external environment and internal environment. The reasonable application of information technology can effectively shorten the time consumed in accounting, and carry out pre-risk control and extended management in the process. Only by taking financial analysis as the basis and providing the corresponding data support, can we avoid the occurrence of financial risk problems. Enterprise financial personnel to clear the requirements of the new accounting standards, and as a basic guidance. Combined with the financial situation of the enterprise, formulate a perfect accounting management system. Enterprises improve the risk warning mechanism by establishing a modern information platform. When there is a big difference between the budget target and the actual situation, the system can give timely warning and take measures to solve the problem as soon as possible. Take an enterprise as an example, the enterprise will financially accounting and management accounting integration, and strengthen the control of cost management, combined with the basic characteristics of product production and quality level. In the production process of different varieties, reduce the expenditure of fixed costs as far as possible, and put forward requirements on production efficiency and quality, and constantly adjust the business situation of the enterprise. Only by controlling the possible risks within a reasonable range can the cost be reduced.

It can be seen from this that management accounting can implement effective financial constraints no matter in the preparation of financial statements or information accounting. Make reasonable use of big data technology to promote the exchange and communication between the financial department and other departments, formulate long-term development plans in combination with the budget objectives and financial conditions of the enterprise, carry out overall forecasting activities in combination with the changes in the market environment, and promote the smooth development of various business activities.

5.3 Establish a Big Data Financial Information Platform and Carry Out Performance Appraisal

First of all, enterprises should optimize the content of financial management and show the advantages of management accounting. Enterprise financial personnel need to analyse the development of big data technology, adjust and optimize the industrial structure. Only by changing the financial content can we realize the data management and innovation of various business projects. In the whole process of financial integration, the external value of the enterprise is analysed, advanced technologies such as big data and cloud platform are scientifically introduced, the work of management accounting is gradually improved, the long-term financial control mechanism is introduced, the investment activities and financial budget of the enterprise are analysed, the management accounting information content is established, and relatively complete accounting management methods are formulated.

Secondly, the establishment of data resource library within the enterprise can not only ensure the accuracy of accounting data, but also provide convenience for financial personnel and management personnel to inquire. By establishing a financial sharing centre, enterprises can
ensure that they can enjoy data information and integrate historical data with the data of various departments. And in a unified platform for accounting, the preparation of sound financial statements, for the major business decisions to provide an important basis. Finally, the work of the enterprise is analysed to realize the combination of big data resource platform and financial data. Carry out comprehensive performance appraisal to improve the comprehensive ability of financial staff. While improving the performance evaluation system, the advantages of management accounting will be brought into play to realize the risk control work and the rational allocation of resources at the same time. As long as the basic management standards are clear and the assessment process of personnel is improved, their work behaviour can be restricted. In the process of the construction of the big data financial management platform, combined with the functions of enterprise business and management accounting, accelerate the speed of the transformation of financial management accounting work, make reasonable use of big data resources, carry out accounting data processing and analysis, so that the efficiency of financial management can be significantly improved.

5.4 Make Rational Use of Big Data

Big data should be applied reasonably in the transition from financial accounting to management accounting. Because big data can improve the integrity and comprehensiveness of management accounting work, management accounting needs to apply big data reasonably and give play to the practical significance of big data background. In order to achieve this goal, Managers can build perfect big data systems. For example, managers can hire technicians to design big data systems and set alarm systems in big data systems. Once there is a security breach in the big data system, the big data system can automatically monitor and issue an alert. Technical personnel can timely understand the cause of this problem, and timely optimize and adjust the treatment, so as to ensure the stable operation of management accounting and achieve the purpose of promoting the transformation of enterprise financial accounting to management accounting [2]. This method can not only improve the financial management level of enterprises, but also further enhance the comprehensive strength of enterprises. In addition, economist Liu Jinzhe pointed out: because the security of this big data system directly affects the stability of enterprise economic operation, in the application of big data, must be reasonable application. If the application of unreasonable phenomenon, it will seriously affect the situation of enterprise financial management. Therefore, in order to solve these problems and ensure the rationality of big data application, special personnel must be found to manage the big data system, regularly check the running status of the big data system as well as loopholes and problems, timely optimize and update the system, and prevent big problems. Only in this way can we improve the operation level of big data systems and enhance the competitiveness of enterprises. In addition, in order to avoid the problem of low application ability of management accounting, it is necessary to build a management team, develop perfect big data application and management standards, and develop application standards. Every management accountant needs to organize information and data according to this standard. Managers in the inspection and management, once found information, data arrangement does not meet the standards of the problem, the managers can be punished in a timely manner, to promote the development of enterprise financial accounting transformation to management accounting under the background of big data.
6  CONCLUSIONS

In a word, the application of management accounting is a huge task, the need for management personnel and each department staff to cooperate, and realize the transformation of traditional accounting to management accounting development. Enterprises need to guide financial personnel, completely change the traditional working concept, and ensure that the management accounting work can meet the needs of modern development of enterprises. Only the integration of financial accounting and management accounting can reduce the consumption of financial resources and human resources, and improve the internal management system. On the basis of the introduction of advanced information platform, the innovation of management accounting concept, ensure the rationality and authenticity of data, promote the economic value of enterprises significantly.

REFERENCES

Design and Implementation of Visualization Application Platform of Enterprise Operation Data Based on R Language

He Zhang
85649133@qq.com
Guangdong Industry Polytechnic, Guangzhou, Guangdong, China

Abstract: This paper studies the design and implementation process of enterprise operation data visualization application platform based on language. In this system, ggplot2 in R language tool is used as the graphic system for making and displaying statistical charts, and javaweb technology is used to develop application programs. R language and Java middleware choose Rserve 1.8.5. Enterprise operation data storage uses big data technology to build hadoop cluster to collect, clean, calculate and save data. In the data classification function module of the system, Bayesian algorithm is used to realize the scientific classification of data. The design of the application system can help the enterprise operators to automatically generate visual reports by selecting the data, which is convenient for users to consult the reference operation data and optimize the operation decisions in time, thus effectively improving the operation quality of the enterprise.

Keywords: R Language, Javaweb, Data Visualization, Big Data, Enterprise Operation.

1 INTRODUCTION

At present, under the background of data explosion caused by the all-round popularization of network equipment, the work of enterprise operation has been inseparable from the evidence support of effective data. This makes enterprises need to continuously increase investment in the collection, analysis and storage of big data. Especially in the operation of enterprises, it is necessary to accurately budget and adjust the operation decisions of enterprises according to the data. Secondly, the person in charge of the operation needs to query and browse the operation data, so as to grasp the overall situation of the enterprise's production and operation, and it is easy to directly attack the weak links of enterprise management to locate the problems, analyze the reasons, realize the improvement of the operation decision, promote the continuous improvement of the enterprise, and form a closed loop of enterprise operation monitoring application. Finally, more market data can be mined through the enterprise operation data, which is convenient for enterprises to find more business opportunities. Data visualization application can help enterprises to provide more efficient guidance. Therefore, this paper studies and designs the visualization application platform of enterprise operation data based on R language. [8]
2 KEY TECHNOLOGIES

2.1 Javaweb

Javaweb refers to the category of application systems developed and designed by applying Java language. Java application uses B/S structure, and its architecture usually consists of four categories, namely servlet or JSP scripting language, HTML page, Java class controlled by business logic and other additional resources. Servlet component is the most representative of servlet interface implementation class. However, JSP components, usually in some specific conditions, are converted into servlet by HTML documents with JAVA programs, such as when clients request JSP files. The definition of Java class refers to the class defined by developers related to web application development programs. Additional resources refer to some static documents and client class files. Static files mostly refer to HTML files stored in the server file system. Client files generally refer to the applet applet and some basic configuration files such as web.xml that run by the client.[1]

2.2 R Language

R is not only a language environment composed of toolkits for data operation, calculation and display, but also has a powerful statistical analysis function. The R language environment contains various statistical analysis and related function models of drawing, which can not only meet the related needs of users' statistical calculation, but also realize the related functions of graphic analysis of data visualization. Besides, R is free and open source software, and it supports multiple platforms, such as Windows, Linux and UNIX. The function of R language can be enhanced according to the expansion package, so as to realize the related functions of quickly manipulating data call, input and output. Users can also use R language to change the source code of the application by expanding packages and plug-ins, so as to realize functions such as extended function loop and self-definition. Through the use of extension packages and plug-ins, R language can even conduct interactive real-time dynamic data analysis through other software, which greatly improves the compatibility and expansibility of R language. [3]

The commonly used charts of data visualization in R language include scatter chart, histogram, histogram, bar chart, box chart, area chart, hotspot chart and relationship chart. The functions used in R language to make charts are ggplot(), geom_histogram, geom_point(), geom_boxplot(), etc. Therefore, R language is used in the visualization application platform of enterprise operation data studied in this paper to process data images and related algorithms. Through R language, intuitive and effective graph list drawing can be automatically generated according to the effective data in the data cluster, and the generated results can be presented to users. The realization process of data visualization is shown in Figure 1. After the data is cleaned and saved by hadoop, the R language and mapreduce are used to construct the drawing function to construct the visual image, which is then presented to the user page of the system through javaweb. [7]
2.3 Rserve

Rserve is an open source data processing middleware. The visualization application platform of enterprise operation data studied in this paper not only uses R language to process and analyze data, but also uses javaweb to deliver data results to users in the form of application page display. Therefore, Rserve is chosen to connect the two technologies. Because Rserve supports TCP/IP protocol of data and network transmission between R language and other languages, it also supports the use of multiple development languages. In this way, the statistical modeling, data analysis and visualization functions of R language can be displayed in javaweb applications through the remote connection of Rserve. [9]

2.4 Development Environment

The development environment of enterprise data visualization application platform is divided into two parts, one is the setting environment of data cluster, the other is the development environment of javaweb application design. The establishment scale of big data cluster, taking into account the data scale of medium-sized enterprises, uses five physical machines to build a big data cluster. In addition to the local management system server of the enterprise, the data comes from a large number of external data of relevant marketing markets obtained by using scrapy and crawler technology. Big data cluster is divided into one primary node and four secondary nodes, and hadoop ecological components are installed on five physical machines to clean, analyze and save big data. The main components of hadoop installed include HDFS distributed storage, mapreduce distributed computing, yarn resource scheduling and hive data warehouse. [6]

The operating system of the physical machine in big data cluster and application design is 64-bit Windows 10, and the CPU is configured as Intel core i7. Rserve, the middleware of R language and Java, chooses version 1.8.5. In the R language tool, ggplot2 is used as the graphic system for making and displaying statistical graphs. The installation package instruction to load ggplot is install.packages("ggplot2"); library (ggplot2). The development language of the application is Java, the development environment is jdk1.7, Apache tomcat 7.54 is selected for the server construction, and MySQL 7.6 is installed on the database server. The framework is SSM framework of spring+springmvc +mybaties. In order to ensure the subsequent development and secondary development of the application system, the relevant configuration of the enterprise operation data visualization application platform is stored in the form of configuration files. The system can obtain the MySQL data connection of JDBC by connecting the database configuration files and reading the key value. The configuration file code of the system is shown in Figure 2.
3 REQUIREMENTS ANALYSIS

3.1 Functional Requirements

The user groups of R-language-based visualization application platform for enterprise operation data are mainly related personnel in charge of enterprise operation. The main functions of the system include data query, data classification, data analysis and data visualization report. Users can make the system automatically generate visual reports by selecting data, which is convenient for users to consult the reference operation data and adjust the operation decision in time. Among them, operation data includes debt repayment ability data, operation ability data, profitability data and financial data analysis. Among them, the operational capacity data mainly includes accounts receivable data and inventory turnover data. According to the relevant industry data captured by big data, the system can also generate a line chart comparing the enterprise with the industry average data. [5]

3.2 Overall Design

As shown in Figure 3, the overall design framework of the enterprise data visualization application platform can be divided into R language data visualization system and javaweb application subsystem. The two subsystems are connected by Rserve middleware, and the communication protocol is TCP/IP. R language subsystem needs to provide data conversion interface and graphics rendering interface. The function of the data conversion interface is to convert the processed data stored and collected by hadoop into a data category recognizable by R language and capable of data processing. And the image rendering interface is used to convert the images generated by R language tools into information forms recognizable by javaweb. The javaweb application part is mainly responsible for the realization of the main functions of the system. This subsystem can be divided into business department and management system. Among them, the business system is a functional module that generates lists and forms for data presentation. The main function of the management system is to realize the permission setting of the system and the setting of user information management module. As a complete application system, its R&D is designed with traditional B/S architecture, so it can't be separated from the technical support of database server. The database server of this system adopts relational database MySQL. [6]
4 FUNCTIONAL IMPLEMENTATION

The main functions of the system include data query, data classification, data analysis and data visualization report. Data classification is a very important link in data processing. This paper uses Bayesian classification algorithm to classify data. The principle of Bayesian algorithm is to first determine that an individual variable is assumed to belong to a random classification, and then calculate the probability that this variable is correct in this classification. Then calculate the probability values of all the classifications in this variable, and compare these probability values to get the best classification of this variable. In the calculation process, we determine the probability value by calculating the ECM misjudgment loss value. The calculation formula of ECM is shown in Formula 1, where $P_1$ and $P_2$ respectively represent the correct probability that a variable belongs to $X_1$ or $X_2$. After that, the minimized ECM value needs to be calculated by Bayesian algorithm. This process is shown in Formula 2, and then the comparison results of the probabilities of all categories are obtained. \(^{[4]}\)

\[
ECM = (R_1, R_2) = L(1/2) \cdot P(1|2) \cdot P_1 + L(1/2) \cdot P(1|2) \cdot P_2
\]  
\[\text{Formula 1}\]

\[
R = \left\{ x \mid \frac{f_1(x)}{f_2(x)} > \frac{L(1|2) \cdot P_2}{L(2|1) \cdot P_1} \right\}
\]  
\[\text{Formula 2}\]

Take Company A as an example. After logging in to the system, the operation leader of Company A can select the data required for data reporting, such as the total market value data of the industry as of August 30, 2022. The system will summarize the collected public data of enterprises in the same industry, and draw the market value data chart of Company A in the same industry. The circular chart version drawn is shown in Figure 4. It can be seen from the figure that the industry valuation of Company A has reached 49.5 billion yuan, reaching the fourth position in the industry, which is higher than the industry average of 10.2 billion yuan.
and the industry median of 387 million yuan. The important index of IPO is the market value. According to the change of market value, operators can judge the overall market positioning of the company on the basis of industry, and then analyze the advantages and disadvantages of the company from multiple dimensions of profit and development, so as to achieve the purpose of deepening the company's operation data.\[10\]

![Figure 4: Ring chart of market value data of company A in the same industry generated by the system](image)

This system mainly evaluates the operation of a company through debt repayment ability $X$ and profitability $Y$ indicators. Taking Company A as an example, the initial characteristic value of debt repayment ability is calculated to be 5.075, with a variance of 49.63%, while the initial characteristic value of profitability reaches 3.859, with a variance of 36.58%, with a total of 86.21%. The variance contribution rate of these two indicators exceeds 85%, which proves that these two ability indicators $X$ and $Y$ can represent the company's operating level. The final evaluation formula is shown in Formula 3.

$$V = \frac{\sigma_1}{\sigma} \cdot X + \frac{\sigma_2}{\sigma} \cdot Y(\sigma = \sigma_1 + \sigma_2)$$  \hspace{1cm} (3)$$

The asset-liability ratio of Company A, which represents the debt repayment ability, is 0.871, the EPS of profitability is 0.824, the net profit score is 0.654, and the return on assets is 0.618. Therefore, the final debt repayment ability is 1.56, and the profitability score is 0.76, among which the net asset income chart of Company A shows the results as shown in Figure 5. What is reflected behind the return on net assets data is the ratio efficiency of shareholders' equity to its own capital, which can often reflect the relationship between a company's profit and the owner's interests. From the chart of return on net assets of Company A generated by R language technology in the system, we can easily see that the return on net assets of Company A has fallen below the average level of the industry, which means that the management has encountered bottlenecks in recent years. Therefore, although Company A's profitability is high, the profit source is mostly the sale of equity, so the overall profitability of Company A is not high, and operators need to pay attention to adjustment. The above data shows that Company A has low profitability but high debt repayment ability, and it still has good operation ability and good development prospects.
5 CONCLUSIONS

This paper provides a good reference for the research of enterprise operation data visualization platform based on R language. However, due to the author's limited ability and time, this paper still lacks more in-depth research and deepening of functions. It is hoped that the follow-up can be improved by relevant experts. The client of the enterprise operation data visualization platform studied in this paper only supports PC browser login, which is slightly insufficient in the current era of popularization of mobile devices. In the future, we will try to study the related content of mobile data processing and further improve the R language operation data visualization algorithm. Nowadays, web crawler technology has developed to the stage where it can generate visual graphics. However, the use of web crawler technology in this paper is limited to text documents, and it needs to be improved.

REFERENCES


Design and Application of Enterprise Intelligent Marketing System Under the Background of Big Data

Benyu Zhao¹,a, Qian Chen¹,b, Yan Hu¹,c*

40684646@qq.com¹, 842134437@qq.com², *295738740@qq.com³

Chongqing College of Architecture and Technology, Chongqing, China¹

Abstract: In order to promote the information economy process of small and medium-sized enterprises in China, promote the prosperity and development of market economy, and improve the problems of poor information management conditions of traditional enterprises and weak internet background marketing ability, this paper establishes an intelligent marketing application software by combining big data technology. The system combines hadoop platform to collect, clean, calculate and process data, and uses javaweb technology to realize data visualization. The FP-growth association rule algorithm can effectively help enterprise managers to understand customers’ needs and consumption preferences, intelligently analyze the most suitable marketing data reports for consumers, and achieve marketing accuracy by means of mass data Internet platform and interactive media, thus further improving the marketing effect.

Keywords: Hadoop, Intelligent Marketing, Big Data, System Design.

1 INTRODUCTION

Nowadays, with the development and popularization of diversified terminals, a large amount of data is generated in the operation of terminal equipment, so the current era is in the era of big data with data explosion. All walks of life need to keep pace with the times and use big data tools to improve their competitiveness. The traditional marketing method is to transmit a large amount of advertising information to network users, but it is easy to lead to the problem of mismatch between advertising information and audience, which brings a lot of negative effects on the marketing products of enterprises and wastes marketing resources to some extent. [10] However, the marketing method using big data technology can collect and sort out a large amount of effective data based on consumers and conduct in-depth analysis, and can draw data reports such as consumer behavior and consumer preferences. Through these intelligent data reports, the marketing executives of enterprises can integrate the consumer groups with relevant commodity attributes, and then establish user groups to realize targeted marketing. According to the above analysis, the author of this paper thinks that the enterprise intelligent marketing system based on big data technology should be developed. The system adopts data mining technology, data warehouse technology and other big data technologies to realize the integration and analysis of business data, and then helps enterprises to make relevant marketing decisions. The development of this system can provide a strong support for the current enterprises to realize information construction. [8]
2 KEY TECHNOLOGIES

2.1 Hadoop Processing Platform

Hadoop platform is the most widely used big data processing platform in the current era, and it has many advantages. Hadoop has good scalability and convenient maintenance; The distributed characteristics make it highly fault-tolerant and highly reliable; Hadoop is a platform made up of multiple components, which has better performance than other platforms. Hadoop was developed in 2004 by Doug Cutting. According to the academic contents of three documents published by Google, its author developed and created them in java language, and chose to publish them in an open source way under Apache software. Hadoop platform is a development and application ecosystem that can support data-intensive applications, and its component team is growing with time. The most important components are distributed file system HDFS and parallel programming model MapReduce. HDFS is responsible for the distributed storage of massive data, while mapreduce is to realize centralized parallel computing of distributed data, and the two complement each other.

As a distributed file storage system, HDFS is a key part of big data technology. HDFS adopts a typical master-slave structure design. The master node names the information file of the slave node, and cuts the information data into multiple blocks, which are saved by different datanode slave nodes. The client sends a data request to the master node namenode, which requests block data from each slave node cluster, and then the slave node sends information to the client to read the returned result data. Hadoop ecosystem is shown in Figure 1. Besides Hadoop and mapreduce, there are many subprojects such as Ambari, Hive, HBase, Zookeeper, Flume and Mahout. With the cooperation of multiple components and clear division of labor, even inexperienced developers can use the advantages of clusters to deal with big data conveniently and quickly. [2]

![Figure 1: Structural diagram of the Hadoop platform](image-url)
2.2 FP-Growth Algorithm

FP-growth algorithm is one of the most classic algorithms in association rule mining. The role of association rule algorithm is to find the relationship between massive data. Compared with other association rule algorithms, FP-growth algorithm has the advantage that it can realize frequent pattern data mining without generating candidate patterns. Next, this article will introduce the operation of FP-growth algorithm. First, we need to traverse the data packet to get the first frequent itemsets, and then we need to traverse the data packet twice and get the support count according to the first data. After that, the records in the database can be sorted and counted. In the sorting process, the infrequent items need to be filtered and cleared away. Finally, the algorithm composition can be iterated according to the mining condition FP-tree to obtain the frequent item set. Suppose we get the frequent itemsets \((f:3,c:4,a:2,b:3,m:4,p:3)\), randomly expand the screening and delete the infrequent items in the data set to establish the memory sorting tree FP-tree. First, select the header table items, take P node as an example, and traverse all the paths involved in P node, so as to obtain \(\{f,a,b,n:3\}\{cp:2\}\) two-day search result paths, which are also the conditional patterns of P. In this random mode, FP-tree mining algorithm and construction algorithm are operated, and frequent itemsets with prefix P can be obtained without generating new paths after continuous iteration.\(^9\)

However, considering that the traditional FP-growth algorithm model has some shortcomings in clear classification, this paper improves the algorithm on this basis. When considering the first frequent itemsets, the category attribute data is no longer taken into account. Instead, the object of counting operation is changed to the relative support number. The implementation process of the improved FP-growth algorithm model is shown in Figure 2.

![Figure 2: Implementation of improved FP-growth algorithm](image-url)
The tree established by the improved algorithm model is named FP-treeR, and the improved algorithm support and confidence formulas are shown in formula (1) and formula (2).\[4\]

\[
support = \frac{supprot\_count + |R\_support\_count|}{2N} \tag{1}
\]

\[
confidence = \frac{supprot\_count + |R\_support\_count|}{2supprot\_count} \tag{2}
\]

2.3 Development environment

In this paper, the author briefly introduces the related technologies of platform development and use. The enterprise intelligent marketing system uses Hadoop as a big data server cluster to process data and store it in MySQL database, and uses JavaWeb technology to develop the corresponding application platform.

According to the data volume and overall operation requirements of the system, this paper chooses to build a Hadoop3.3.1 cluster with three nodes, including a master node and two slave nodes. The main node is named namenode, and the secondary node is named datanode. Then, the distributed collaboration system zookeeper-3.4.1, distributed file system HDFS 2.6.5, flume1.9.0, Hive 0.13.1 and Hbase2.6.5 are installed and deployed in these three nodes synchronously, and the initial construction of hadoop cluster is completed. The cluster will be developed under Linux system. This paper selects Centos6.5 Server release version of Linux operating system.

The front-end development tool used in the JavaWeb application of this system is bootstrap+jquery, and the development language is JavaScript+HTML+CSS. The back-end Java development tool is IDEA 2021.1.3 (Ultimate Edition), the development environment is JDK 1.8, and the J2EE framework of Tomcat+Spring MVC+Spring+MyBatis is is used in the implementation of this system. The development language is Java, and MySQL 8.0.28 is selected to help manage data. Through the introduction of the above key technical theories, the overall environment of enterprise intelligent marketing system development, the configuration of related software and tools are determined, and the technical feasibility of the overall project is also clarified.\[3\]

3 REQUIREMENTS ANALYSIS

3.1 Functional requirements

The user group of the intelligent marketing system of enterprises is positioned in the marketing managers of enterprises. Firstly, the system should acquire and integrate the marketing data of the enterprise management system, and save these data in the system cluster, so it has the functions of data warehouse, search engine and classification. Secondly, the system should perform algorithm analysis on the data, and the system should have the function of issuing data analysis reports.

Finally, the data analysis report should have the function of data visualization in order to be more intuitive and concrete, and make it easier for users to understand. The data index system
in data warehouse includes customer attribute, customer consumption attribute and customer gender. The daily table update data of the system includes daily increment data and daily total data. Data labels include statistical labels, rule labels, mining labels, calculation labels, etc. The analysis includes commodity analysis, user analysis, traffic analysis and recommendation effect analysis. \(^{(1)}\)

3.2 Overall design

Enterprise intelligent marketing system adopts B/S structure and is designed with multi-layer distributed idea. The overall design architecture diagram of the system is shown in Figure 3. The system is divided into customer layer, web layer, business layer and database layer. The client layer includes pages and application clients, and users access the system through PC user ports. The main function of Web layer is to collect and process requests from clients. The content of web layer includes javabean objects and JSP pages. The main function of the business layer is to handle business logic and realize the interface function between the client and the system server, and encapsulate a large number of functional components. The data layer is responsible for managing the database server, which is convenient for the system to immediately call the data resources. \(^{(5)}\)

![Figure 3: Overall architecture diagram of the system design](image)

4 FUNCTIONAL IMPLEMENTATION

The main functions of enterprise intelligent marketing system can be divided into data warehouse, data classification, data query, data analysis and visual report. This section introduces the implementation logic of data analysis report viewing. When the user clicks on the front page of the data analysis function, selects the data range, and then clicks "Start Data Analysis" to start generating the data report. The back-end server will use the SummaryController to receive the get request sent by the front-end, and obtain all relevant data in MySQL, the database server, through the getDigest method. \(^{(6)}\)

In order to improve the security of enterprise management information system, the system needs to judge the legality and permission category of system users. The user names and IP addresses of all relevant users will be saved in the background of the system database. Only those who meet the requirements of the database can log in normally. In the data retrieval function module, the user presentation layer of the system will display the retrieval conditions
related to the information categories queried by users, and the search engine of the system will make retrieval calls according to the classified keywords. Data access uses the encapsulated business data control component in the system, and after the results are found, they are returned to the presentation layer of the browser and fed back to the users. Among the data warehouse modules, two types of input modes are supported, and the automatic input mode supports the information input of other department information management systems of enterprises. The manual input method supports the salesman to manually input all kinds of sales order data of the enterprise into the background database of the system. The system will automatically enter the input information into the background database and establish statistical tables.

In this paper, the improved FP-growth algorithm is used to analyze the intelligent marketing data. The operation process is to first calculate the frequent data set of the project, then determine the association rules between consumers and customers to generate customer interest groups, and then get the recommendation rules of marketing products according to the preference list data of this group. The performance test is required for the intelligent marketing data analysis function of P-growth algorithm. Taking the marketing database of Company A as an example, we first select four data sets for functional testing. The data sets are shown in Table 1. [7]

<table>
<thead>
<tr>
<th>Data set</th>
<th>Transaction item quantity</th>
<th>Data capacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2000</td>
<td>30MB</td>
</tr>
<tr>
<td>2</td>
<td>5000</td>
<td>45MB</td>
</tr>
<tr>
<td>3</td>
<td>8000</td>
<td>100MB</td>
</tr>
<tr>
<td>4</td>
<td>15000</td>
<td>150MB</td>
</tr>
<tr>
<td>5</td>
<td>20000</td>
<td>200MB</td>
</tr>
<tr>
<td>6</td>
<td>80000</td>
<td>700MB</td>
</tr>
</tbody>
</table>

After that, we use Hadoop cluster to use the improved FP-growth algorithm to process data and control the same support variables. The data is processed in descending order of data capacity, and the support level is set to 15%. The data processing results are shown in Figure 4. By observing the data in the figure, we can see that the running time of the improved FP-growth algorithm studied in this paper is less than that of the traditional algorithm, thus achieving the purpose of improving the data processing efficiency.
5 CONCLUSIONS

This paper studies the intelligent enterprise marketing system based on big data technology. The FP-growth association rule algorithm of this system can effectively help enterprise managers understand customers' needs and consumption preferences, and intelligently analyze the most suitable marketing data reports for consumers. Therefore, this research has certain application value. However, due to the author's limited ability and time, there are still some shortcomings to be improved. The intelligent marketing system designed in this paper can assist marketing leaders to make relevant decisions, but it can't directly realize the automatic intelligent marketing of the Internet platform, so there is still some room for development in intelligent marketing decision-making.

REFERENCES

Value Co-Creation of Cultural Tourism Enterprises' Service Ecosystem in Digital Transformation: -- Case Study of Digital Shaanxi Tourism Group

Lina Sui¹, Kailun Guan¹,a
Email address: guankailun@163.com

School of Tourism & Research Institute of Human Geography, Xi’an International Studies University, Xi’an 710128, China¹

Abstract: This study focuses on how cultural tourism enterprises promote value co-creation in service ecosystems with the help of digital technology, taking digital Shaanxi tourism as the research object for case analysis, and exploring the evolution mechanism of value co-creation in tourism service ecosystem under the context of digital transformation by using the grounded method. It is found that digital technology plays an important facilitating role throughout the value co-creation process, coordinates the layout of multiple subjects in the system, and promotes resource sharing, which can improve the frequency and precision of demand communication and achieve cost reduction and efficiency. The market is a key factor influencing the value co-creation of the service ecosystem of cultural tourism enterprises, and plays a role of precision, coordination and adaptability from the bottom up, driving the integration of customer value, platform value and system value. This paper provides new theoretical insight into the evolutionary mechanism of value co-creation of cultural tourism enterprises' service ecosystem in digital transformation and provides an important practical reference for cultural tourism enterprises to improve their viability and strengthen their resilience.

Keywords: Digital transformation, Cultural tourism enterprise service ecosystem, Value co-creation, Shaanxi Tourism Group.

1 INTRODUCTION

With the gradual spread of a new generation of information technologies represented by SMACIT (the convergence of social and mobile media, analytics, cloud and IOT), inter-organizational data sharing, scenario-based instant applications, and large-scale data applications have become possible. The technology application of cultural tourism enterprises has thus evolved from mere information transfer to value transfer and value creation. However, the digital transformation of cultural tourism enterprises generally faces challenges such as weak technology transformation ability, the insufficient resonance between digital technology and demand, deficient innovation in publicity and communication, and lack of compound talents with high digital literacy.

In recent years, the research field of value co-creation is no longer limited to binary and network relationship perspectives and the multi-interest subject's diversified interactions in the
service ecosystem perspective have become the focus of scholars and business practices [18], and researchers have begun to pay attention to the systematic value creation orientation among multi-participants, proposing that value co-creation embedded in micro, meso and macro structure levels is a resource integration process [9]. Existing studies on value co-creation in service ecosystems have mainly conducted useful discussions in terms of interaction structures and facilitation conditions [22], but the impact of social environment and tourism demand transformation on tourism value co-creation is underappreciated.

The development from simple value interactions to complex and diversified service ecosystems in the tourism field is the inevitable path, and future competition in the cultural tourism market will further evolve into a competition between ecosystems. However, most of the existing studies are reviews based on theoretical deduction, while empirical studies based on enterprise practice are still lagging. Given this, this paper adopts the case study approach and selects Shaanxi Tourism Group as a case study object to explore how the ever-emerging and evolving digital intelligence technology can promote the value co-creation and evolution of the cultural tourism enterprise service ecosystem at various levels and stages, to realize the activation and enhancement of its viability. To answer the above questions, this paper constructs a theoretical analysis framework of “Value proposition, Value network reconstruction and Co-creation value realization”, deconstructs the evolution mechanism of value co-creation in the service ecosystem of cultural tourism enterprises in the digital context, and compensates for the existing literature to a certain extent.

2 LITERATURE REVIEW

2.1 Digital transformation of cultural tourism enterprises

In the early stage, the digital transformation application of cultural tourism enterprises mainly relied on the Internet to provide tourism information and book tourism products and services online. However, in the new era, cultural tourism enterprises are more concerned about the agility of responding to market demands and the intellectual support of enterprises adapting to market development, promoting the digital trend of production [12], experience [1] and service mode [23]. The quality and efficiency of services provided by cultural tourism enterprises will be improved and new vitality of the cultural tourism industry will be stimulated through new forms and models such as product intelligence [10], marketing precision [19], personalized demand [17] and online enterprise services [4].

At the same time, the COVID-19 pandemic has also forced cultural tourism companies to accelerate their online and digital transformation and upgrade to adapt to market changes. However, in general, the current research on the digital transformation of cultural tourism enterprises is mostly limited to performance improvement and business expansion, and the research on the business model and organizational structure reform needs to be in-depth, and there is a lack of empirical research rooted in cultural tourism enterprises.

2.2 Value co-creation studies in service ecosystems

The value co-creation in service ecosystems includes the dynamic evolution of micro, meso and macro levels [3], focusing mostly on the multi-level interactive structure [21], regime and
market roles in value co-creation, where diverse participants agents are connected through direct or indirect relationships, and the same value proposition becomes the basis for participants to establish new connections [11] and carry out value creation to achieve the desired outputs [14]. In the context of digitalization, digital technology is an important driving force for value co-creation, as it reshapes the value co-creation of innovative subjects in the service ecosystem, increasingly participating in interactive value creation as non-human participants, and converging previously dispersed demands and experiences across spatial boundaries to digital platforms [13]. However, few existing studies have discussed the changes in the value co-creation formation mechanism of the service ecosystem based on enterprise perspective in the digital context, and the multi-agent composition and interaction among the agents need to be further explained.

2.3 Value co-creation in tourism

The concept of service-oriented logic provides a new direction for tourism development, with the realization of the important impact of co-creation and even co-destruction brought by information technology, studies have explored social media interactions [16], co-creation behaviors [15], and tourism strategies developed with the participation of web users [7]. For cultural tourism enterprises, promoting value co-creation requires building a co-creation platform, improving its ease of use, aesthetic experience, and trust and thus influencing tourists’ fit behavior [22]; identifying value-added links for customer participation at each stage of the value chain and designing co-creation approaches and contents [8], but the value co-creation process in the field of tourism is a black box that needs to be continuously explored, and there is currently insufficient attention to the position and role of the cultural tourism enterprise in the value co-creation process, the impact of shifts in the social environment and tourism demand characteristics on tourism value co-creation has not been fully paid attention to.

3 RESEARCH DESIGN

3.1 Research methods and case selection

This paper adopts the case study method [20], selects cases through the theoretical sampling method [6], and takes into account the availability and typicality of case data. And the grounded theory and coding analysis methods [2] are used to conduct an in-depth analysis of the research process. The cooperation among related enterprises in the Digital Shaanxi Tourism network has typical dynamic evolution characteristics, which meet the research conditions of the service ecosystem.

Shaanxi Tourism Group is one of China’s earliest cultural tourism groups to practice digital transformation, with a complete industrial chain, prominent comprehensive effect, and extensive influence. In 2020, Shaanxi Tourism Group built the industry's first "Digital Cultural Tourism Lab", which became the incubation base for digital projects. In early 2022, Shaanxi Yuanjing Digital Innovation Technology Co., LTD. (hereinafter referred to as "Yuanjing Digital Innovation") was established to further explore the application scenarios and build the industrial ecology. In December 2022, Shaanxi Tourism Group was awarded "Top 20 Chinese Tourism Groups" for three consecutive years. At present, almost all mainstream technology
and media support enterprises have cooperation with Shaanxi Tourism.

<table>
<thead>
<tr>
<th>Table 1: Case study business information</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main business area</strong></td>
</tr>
<tr>
<td>Digital cultural tourism creative content application services</td>
</tr>
<tr>
<td>Tourism complexes/scenic spots</td>
</tr>
<tr>
<td>Tourism Services</td>
</tr>
<tr>
<td>Technical support</td>
</tr>
<tr>
<td>Media support</td>
</tr>
<tr>
<td>Channel Support</td>
</tr>
<tr>
<td>Market Positioning Support</td>
</tr>
<tr>
<td>Content Creation Support</td>
</tr>
</tbody>
</table>

3.2 Data sources and data collection

This paper combines primary and secondary data and uses triangulation to ensure that the research data are interactively supplemented and verified by multiple information sources. The sources and quantities of data are shown in Table 2.

<table>
<thead>
<tr>
<th>Table 2: Case data sources</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data type (code)</strong></td>
</tr>
<tr>
<td>Interview transcripts (D1)</td>
</tr>
<tr>
<td>Field observation records (D2)</td>
</tr>
<tr>
<td>Website News Reports (S1)</td>
</tr>
<tr>
<td>Conference Materials (S3)</td>
</tr>
<tr>
<td>Core Journal Literatures (S4)</td>
</tr>
</tbody>
</table>
3.3 Data coding and analysis

3.3.1 Open coding

This study obtained 425 items of secondary data. Through open coding, 51 initial concept labels (N1-N51) were obtained, including 42 high-frequency concepts (HN1-HN42). Combining similarity concepts yields 28 subcategories. Table 3 shows examples of open coding at different levels in this study.

Table 3: Examples of opening coding

<table>
<thead>
<tr>
<th>Examples of items</th>
<th>Conceptual construction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Huaqing Palace combines AR, visual intelligence algorithms, and other technologies, creating the scenery mentioned in the poem, allowing visitors to have a &quot;dream back to the Tang Dynasty&quot; participation experience.</td>
<td>HN1 Storytelling</td>
</tr>
<tr>
<td>Yongcn Hotel started to analyze the needs of different regions and types of guests from online booking, and actively implemented and continuously tracked them. The &quot;Shaaxi Travel high GO&quot; live broadcast room focuses on Shaanxi Tourism's special resources and local special products, providing visitors with more intuitive and preferential access to product information and enabling consumers to participate in activities.</td>
<td>HN3 Living service</td>
</tr>
<tr>
<td>Huaqing Palace has explored a &quot;micro-innovation to patent&quot; value cashing model to transform quality achievements into brand assets and converts the innovation technology accumulated in the process of tourism quality management into brand assets.</td>
<td>HN5 Live interact</td>
</tr>
<tr>
<td>The successful holding of the 14th Games has brought a strong brand radiation effect to the Shaanxi International Sports Window, and the surrounding business district and sports clusters have accumulated a great atmosphere and resources.</td>
<td>HN25 Copyright manage</td>
</tr>
<tr>
<td>Since the birth of Hi Metaverse, it has insisted on technology empowerment and is committed to excavating the spiritual connotation of cultural resources, making the platform an important carrier for the general public to enhance cultural self-confidence.</td>
<td>HN27 Leveraging</td>
</tr>
</tbody>
</table>

3.3.2 Spindle coding

By analyzing the organic correlation of the meanings of 28 sub-categories, this paper further gets the corresponding main categories (Table 4).

Table 4: Examples of spindle coding

<table>
<thead>
<tr>
<th>Main category</th>
<th>Level</th>
<th>Associated sub-category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value proposition</td>
<td>1</td>
<td>A1a In-depth experience;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A1b Interactive marketing;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A1c Fine service;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A1d Intelligent management</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>A2a Partner matching;</td>
</tr>
<tr>
<td></td>
<td>meso</td>
<td>A2b Industry identification;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A2c Enterprise vision</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>A3a Opportunity recognition;</td>
</tr>
<tr>
<td></td>
<td>macro</td>
<td>A3b Mission commitment;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A3c Demand of the Times</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B1a Demand communication;</td>
</tr>
<tr>
<td>Value network</td>
<td>1</td>
<td>B1b Demand matching;</td>
</tr>
<tr>
<td>reconstruction</td>
<td></td>
<td>B1c Resource optimization</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>B2a Resource reproduction;</td>
</tr>
<tr>
<td></td>
<td>meso</td>
<td>B2b Resource integration</td>
</tr>
</tbody>
</table>
3.3.3 Selective coding

By continuing to examine the 9 main categories and 27 associated sub-categories, repeatedly comparing the original data records, the core category of "value co-creation of cultural tourism enterprises’ service ecosystem in digital transformation" was extracted, and through matching, interaction and fusion, the precision, coordination and adaptability of cultural tourism enterprises were improved, the value of the customer, platform and system is realized, and the next round of value co-creation activities was paved and supported.

<table>
<thead>
<tr>
<th>Level</th>
<th>Canonical relation</th>
<th>Key phenomenon</th>
</tr>
</thead>
<tbody>
<tr>
<td>Micro</td>
<td>Proposition</td>
<td>Novel and intelligent sensory experience brings more efficient and intimate fine service.</td>
</tr>
<tr>
<td></td>
<td>Precision</td>
<td>Use digital technology to precisely communicate and match enterprises, employees and consumers.</td>
</tr>
<tr>
<td></td>
<td>Reorganization</td>
<td>Use data resources to meet customers' customized and diversified digital consumption needs.</td>
</tr>
<tr>
<td></td>
<td>Matching</td>
<td>Establish a co-creation platform to further promote industry chain resources and business synergy.</td>
</tr>
<tr>
<td></td>
<td>Realization</td>
<td>Revitalize idle resources, integrate online and offline traffic, and realize data interoperability.</td>
</tr>
<tr>
<td></td>
<td>Customer value</td>
<td>Create a standardized system to promote cooperation and symbiosis, and support intelligent decision-making.</td>
</tr>
<tr>
<td></td>
<td>Proposition</td>
<td>Leverage the situation by identifying opportunities and proactively responding to the development of the needs of the times.</td>
</tr>
<tr>
<td>Meso</td>
<td>Reorganization</td>
<td>Integrate the elements of system capacity, make overall plans, and fusion intelligent business.</td>
</tr>
<tr>
<td></td>
<td>Interaction</td>
<td>Through supply-side structural reform, demand-side management maximizes benefits.</td>
</tr>
</tbody>
</table>
4 CASE ANALYSIS

This paper adopts the case study method, focuses on the constantly upgrading technologies to empower the service ecosystem of cultural tourism enterprises, and divides the service ecosystem with Shaanxi Tourism Group as the core into three phases, which can be successful because it transforms the original operation mode, continuously innovates to create value propositions, emphasizes the depth of experience and service refinement, and coordinates the participation and interaction of multiple subjects, especially heterogeneous subjects, drives the synergy, sharing and reconstruction of resource elements in the service ecosystem of cultural tourism enterprises, and achieve the win-win and all-win of multiple subjects. This section will deeply analyze the construction process of the model and elaborate on the evolution of value co-creation of Shaanxi Tourism Group's service ecosystem under digital transformation.

![Figure 1. Value co-creation evolution process of cultural tourism enterprise service ecosystem in digital transformation](image)

4.1 Phase I: Exploration

To meet the in-depth experience needs of indigenous people in the digital era, the value proposition of "innovative service scenarios" is proposed, focusing on immersive experience business and cultivating market-recognized products and services. In the context of a recurrent epidemic, Shaanxi Tourism also actively carries out self-rescue by trying to develop "small but beautiful" projects, and the scope of application has gradually spread from the initial performances to exhibitions, secret rooms, live broadcasts, blocks, amusement parks, and other scenes.

Since 2016, from "Northern Expedition Memorial" and "Red Detachment of Women" to "Night Fantasia of Bailuyuan" and "Hi Metaverse", Shaanxi Tourism Group has seized market opportunities, deeply perceived the real needs of tourists, continuously used emerging technologies to fully mobilize tourists' senses, satisfied or even created tourists' experience
needs, effectively extended tourists' stay time by changing their consumption patterns, and obtained the recognition of young tourists in terms of values and lifestyle through in-depth interaction, thus created new value for participants within the system. In the context of a recurrent epidemic, Shaanxi Tourism also actively carries out self-rescue by trying to develop "small but beautiful" projects, and the scope of application has gradually spread from the initial performances to exhibitions, secret rooms, live broadcasts, blocks, amusement parks, and other scenes.

Figure 2. The first phase of the Value co-creation evolution process

However, due to the imperfect cognition of digital cultural tourism on the demand side, the limited application scenarios of digital technology, the immature management and operation practices of the supply side, and the monitoring of effects have not yet formed a widely recognized evaluation, Shaanxi Tourism Group's overall organizational efficiency and management mode also need to be rapidly upgraded in the digital era, and the core participants promote the cultural tourism service ecosystem to the second stage in the exploration aimed at solving the above issues.

4.2 Phase II

In the context of the epidemic, the demand of cultural tourism enterprises to reduce operation and maintenance costs has become increasingly urgent, and barriers between business units have become the biggest obstacle to the utilization of data resources. Therefore, Shaanxi Tourism Group puts forward the value proposition of "integrating data, controlling costs, and increasing efficiency", mines and applies data, breaks business silos, and promotes the integration of online and offline processes, making the system data interworking while focusing on improving data security strategies to ensure data security. With the accumulation of data assets, Shaanxi Tourism Group has a broader range of data fusion, forming a data interface to connect with external partners, and establishing a data fusion and sharing mechanism.
Since 2020, Shaanxi Tourism developed a big data visualization platform, launched Pocket Shaanxi Travel APPs and applets, realized office automation, online business, data standardization, and intelligent management, improved organizational efficiency and internal collaboration capability, strengthened project supervision, to respond the needs of the market, customers and employees timely. However, long-term development requires sustainable competitive advantages through innovative business models and brand appreciation, which places higher demands on marketing accuracy, market agility, and adaptability to the times. How to gain insight into the key trends of the cultural tourism industry, deliver information more efficiently and accurately and implement optimal strategies? It has become the driving force for Shaanxi Tourism Group to promote the next iteration of the service ecosystem.

4.3 Phase III

Shaanxi Tourism Group’s media center provides one-stop analysis and operation, insight into market demand, intelligent scheduling, fine delivery, and real-time interaction with online consumers, puts forward the value proposition of "brand appreciation and long-term development", and strengthens resilience capacity building to maximize benefits. In the third phase of the value network reconstruction process, the accumulation of data assets in the earlier stage has formed a digital product matrix, generating scale effects, and building a 1+N ecosystem.
In 2021, Shaanxi Tourism Group together with Ant Technology Group established the first national cultural tourism digital industry platform, launched the "Cloud Shaanxi Travel" super APP, and opened the "Shaanxi Travel High GO" live studio, which attracted widespread attention. In addition, Huaqing Palace Scenic Area of Shaanxi Tourism won the "2021 National Copyright Demonstration Unit" also represents that the group's awareness of intellectual property has been further strengthened, which will help to improve the trust of all participants in the system, stimulate the enthusiasm for value co-creation, establish a long-term mechanism, and form a stable mutually beneficial and symbiotic relationship. However, there are also some hidden questions, such as: Does intelligent technology violate the privacy rights of tourists? Are cultural tourism enterprises able to support changes and strengthen system implementation capabilities in the long term? Based on theoretical logic and existing practice, it is believed that blockchain technology and the cultivation of high-quality digital cultural tourism talents will become the enabling power for the next stage. The blockchain-based integration platform can effectively enhance the trust, security and transparency of the cultural tourism industry. Meanwhile, it is also necessary to establish a talent guarantee mechanism and perfect the digital cultural tourism talent training system to help achieve high-quality and sustainable development of the cultural tourism industry.

5 CONCLUSION AND PROSPECT

5.1 Conclusions

Based on the case analysis of Shaanxi Tourism Group, this study extracts the evolutionary mechanism of value co-creation of the cultural tourism enterprise service ecosystem in digital transformation (Figure 2). Based on this model, the following main conclusions are drawn.
Digital technology, as a facilitator, supports the value co-creation of the whole process at all levels of the cultural tourism enterprise service ecosystem. At the macro level, digital technology connects multiple subjects including government agencies, industry associations, universities and colleges, media and the public, etc., and each participant maximizes the benefits through the overall layout. At the meso level, digital technology plays a role in connectivity, transcends the traditional boundary between platforms and enterprises, integrates assets, management, industrial chain and other capability elements, and promotes the integration and symbiosis of online and offline. At the micro level, digital technology transforms data into resources, improving the breadth and precision of demand matching among cultural and tourism enterprises, employees and consumers. In the digital environment, digital technology connects scattered subjects, demands, resources and data, coordinates the participation and interaction of multiple subjects, especially heterogeneous subjects within the system, promotes cost reduction and efficiency increase, and drives the coordination, sharing and reconstruction of resource elements within the service ecosystem of cultural tourism enterprises.

In the context of digital transformation, the market is an important promoting factor for value co-creation in the service ecosystem of cultural tourism enterprises. At the micro level, based on customer-oriented demand communication and matching, cultural tourism enterprises create new consumer demand and usher in new market opportunities. At the meso level, cultural tourism enterprises do a good job in partner selection and resource coordination, establish an integrated platform based on the principle of the open system, and realize the enhancement of platform value while transforming the cultural tourism business model. At the macro level, cultural tourism enterprises identify important opportunities to comply with the market development trend and drive the digital upgrading of the cultural tourism consumption market. The market breaks through the micro, meso and macro level boundaries to play a precise, coordinated and adaptive role respectively, gradually realizing the integration of customer value, platform value and system value, agilely responding to market demand, and ultimately strengthening the resilience of cultural tourism enterprises.

5.2 Practical Inspiration

At present, the impact of COVID-19 on China's society and economy is still ongoing. Exploring the constitutive dimension and evolutionary mechanism of multi-level value co-creation in the service ecosystem of cultural tourism enterprises is helpful to clarify the vague understanding of value creation in the complex system structure, dissect the changes of key elements such as subjects, relationships and resources in the system, and provide a practical reference for decision-makers to promote digital transformation.

(1) Strengthen the top-level design, improve the agile response-ability to the market, and deeply grasp the development trend of "diversification, interaction, precision, quality" of the cultural tourism market. (2) Attach importance to the coordinating role of the cultural tourism enterprise service ecosystem platform, stimulate all kinds of innovation subjects to actively integrate into the ecosystem, and allocate innovation resources reasonably and efficiently. (3) Attach importance to the transmission role of digital technology. Through the intelligent interconnection of data, the adaptive digital enabling mechanism is introduced to assist scientific decision-making, and digital technology will strengthen the multi-sensory enjoyment of cultural travel experience, and enhance the interest and attraction of cultural travel products.
5.3 Research limitations and outlook

This paper focuses on the evolutionary mechanism of value co-creation in Shaanxi Tourism Group up to 2022, and future research can lengthen the time dimension to discuss the dynamic evolutionary mechanism of value co-creation in the service ecosystem from a longer-term perspective. Furthermore, the practice of change based on emerging digital intelligence technologies is generally in the exploration stage, in the future, it is still necessary to track new innovative case practices, and multi-case analysis can be carried out based on the development practices of other advanced cultural tourism groups, and the mechanism of deep integration of cultural tourism industry and digital economy can be explored from the aspects of organizational structure reform and external utility verification.

REFERENCES

The Correlation Analysis Between the Informatization Level and the Profit Growth Rate of Small and Medium-Sized Enterprises

Xingcheng Guo*

*Corresponding Author: guoxc2@gcu.edu.cn

Guangzhou City University of Technology, Guangzhou, China

Abstract: In recent years, the important field of information management is to improve the informatization level of Small and Medium-Sized Enterprises (SMEs). Background: At present, with the continuous development of computer technology and the gradual realization of informatization in all walks of life, the application of information technology in small and medium-sized enterprises has a direct impact on work efficiency and profit growth. Methods: This paper investigates 6 small and medium-sized enterprises (SMEs) to empirically studies the positive relationship between enterprise informatization level and enterprise profit growth rate by using the gray correlation analysis method. Results: The level of informatization is positively related to the growth rate of interest rates in small and medium-sized enterprises (SMEs). Conclusion: SMEs should optimize the construction of enterprise information infrastructure platforms, promote the standardization and integration of enterprise information systems, create a sound information data network system, and strengthen the training of enterprise informatization talents in the future.

Keywords: Small and Medium-sized Enterprises (SMEs), Informatization level, Grey Correlation Analysis Method.

1 INTRODUCTION

Nowadays, the information age has come, and people pay more attention to information technology, and gradually devote themselves to the information industry [5]. In small and medium-sized enterprises (SMEs), information technology is very critical, and relevant information personnel needs to pay attention to it. However, at this stage, there are still some small and medium-sized enterprises (SMEs), whose economic development is limited by information technology, because the development level of these enterprises is not high, and at the same time, the internal facilities of these enterprises are not perfect, making them unable to keep up with the pace of the times. At present, SMEs focus on the profit growth rate of enterprises. Is there a certain correlation between the profit growth rate of SMEs and their information level? If the level of enterprise informatization level is positively related to the growth of enterprise profit margin, the enthusiasm of SMEs for information resource management will be improved, and the amount of investment in all aspects of information resource management will also be increased, which will greatly promote the healthy and sustainable development of SMEs.
2 LITERATURE REVIEW

Under the background of the information age and knowledge economy, the problem of information management has become an unavoidable problem for small and medium-sized enterprises, which account for a significant proportion of the organizational form in the economic structure, and an important fundamental problem related to the competitiveness and operation of small and medium-sized enterprises and even the economic development of each country. With the development of the market economy, SMEs are under increasing pressure. To improve their competitiveness, SMEs must upgrade their management, strengthen information construction and implement information management. Many small and medium-sized enterprises have already started their information management plans, but many small and medium-sized enterprises have many problems that cannot be ignored in information management. The adoption of information management will help promote the innovation and development of enterprises in all aspects of the operation process. Enterprises will operate more smoothly, help enterprises save human costs, and maximize the use of enterprise resources to help enterprises save resource costs. Due to the small scale and limited capital and other resources of small and medium-sized enterprises in China, the level of enterprise informatization development is low. At present, most small and medium-sized enterprises in China are still in the primary stage of informatization development. Many SMEs are still willing to choose traditional methods such as manual recording in information management, human resource management, accounting information management, etc. Compared with large domestic enterprises, the informatization development of small and medium-sized enterprises is far behind. For enterprise operations, the goal is to improve productivity, and performance evaluation is an important organizational form and management means to define goals. Therefore, in enterprise informatization construction and information system application, evaluating the performance of informatization application is an important management means to achieve the goal of informatization application. Only through performance evaluation can we clearly understand the efficiency of enterprise informatization input and output. Some managers have divorced their understanding of informatization from reality and invested large costs in building management information systems. However, the enterprise informatization systems built are inconsistent with the management business processes, leading to idle resources. Enterprise managers should have a clear understanding of the role of information technology. According to relevant data, currently, small and medium-sized enterprises have not formulated corresponding institutional measures in the development of implementing information management mode, which makes the information management mode unable to play its maximum role in time. The construction of information management requires not only the purchase of hardware equipment but also the purchase of software equipment, maintenance fees, personnel training fees, and other related expenditures. The overall cost of information construction is not a small amount. Many small and medium-sized enterprises do not have the strength to support it, and information construction is a continuous investment. Therefore, the fund shortage of SMEs has a certain impact on the efficiency and efficiency of enterprises, which is also an important reason why the informatization construction of SMEs cannot keep pace with the development of the times.

Enterprise informatization is people-oriented. It is not simply computer matching, but more importantly human factors. It is necessary to recognize what enterprise informatization is and how carry it out. How to promote the faster development of enterprises with the help of enterprise
informatization is the key to the construction of enterprise informatization. The active participation of employees will have an impact on the effective operation of the whole system. Enterprises must be based on their talent exploration and training, people-oriented, all staff training, establish a good enterprise information team, and promote enterprise information management \[^{13}\]. In addition to the impact on the development of enterprises, the level of informatization will also affect the demand for enterprise elements. A supply chain is an important way for enterprise informatization to play the above role. The improvement of enterprise informatization reduces the information cost of both supply and demand sides in the financing market, which is conducive to the development of virtual economy supply chain finance. \[^{4}\]. Relying on the Informatization Association, Internet Association, and various industry alliances, we will hold enterprise informatization expert seminars, Internet on-site meetings, in-depth integration of "industrialization and industrialization" and other activities to awaken the demanded awareness of all demand subjects for enterprise informatization construction, change their ideas, and create a good atmosphere for traditional industries, small and medium-sized enterprises to achieve transformation and upgrading through cloud computing and industrial Internet to accelerate informatization. Create an urgent awareness of information construction in the whole society, especially among senior managers of enterprises, improve the information awareness of cadres, masses, and enterprise employees, understand information construction knowledge and create an atmosphere of information knowledge sharing and continuous learning. \[^{8}\].

3 MATERIALS AND METHODS

3.1 Research Object

This time, the author investigated 6 SMEs in Guangzhou Shawi SME Park and mastered various relevant data on their informatization. According to the requirements of the enterprise, the name of the enterprise is replaced by A.B.C…. These six enterprises label them as (1) Guangzhou A Solid Wood Furniture Co., Ltd., (2) Guangzhou B Technology Co., Ltd., (3) Guangzhou C Industry Co., Ltd., (4) Guangzhou D Plastic Technology Co., Ltd., (5) E Electronics Industry Co., Ltd., and (6) F Shoes Co., Ltd. According to the results of the questionnaire, the order of profit growth rate of the six enterprises is (5) (2) (4) (6) (3) (1). The following is an empirical analysis using the correlation analysis method in the gray system theory. The following is an empirical analysis of the correlation between the informatization level of these six SMEs and their profit growth rate using the gray correlation analysis method.

3.2 Research Method

The basic analysis steps of the grey relational analysis method are as follows:

(1) Establish evaluation indicators, collect relevant data, and establish a matrix consisting of n data sequences (1):

\[
(X'_1, X'_2, \ldots, X'_m) = \begin{pmatrix}
X'_1(1) & X'_2(1) & \cdots & X'_n(1) \\
X'_1(2) & X'_2(2) & \cdots & X'_n(2) \\
\vdots & \vdots & \ddots & \vdots \\
X'_1(m) & X'_2(m) & \cdots & X'_n(m)
\end{pmatrix}
\]
The maximum value of each index is used to establish a reference number series, and the original data matrix is standardized.

The absolute difference between the index sequence (comparison sequence) of each evaluated object and the corresponding element of the reference sequence is calculated in turn. That is \(|X_0(k) - X(k)|\), which \(k=1,2,3,... \ n\); \(i=1,2,3,... \ n\) (\(n\) is the number of evaluation objects).

(4) Determine \(\min_{i=1}^{n} \min_{k=1}^{m} |X_0(K) - X_i(K)|\) and \(\max_{i=1}^{n} \max_{k=1}^{m} |X_0(K) - X_i(K)|\)

(5) Calculate the correlation coefficient. The correlation coefficient between each comparison sequence and the corresponding element is calculated by the following formula:

\[
\rho = \frac{\min_{i=1}^{n} \min_{k=1}^{m} |X_0(K) - X_i(K)| + \rho \cdot \max_{i=1}^{n} \max_{k=1}^{m} |X_0(K) - X_i(K)|}{|X_0(K) - X_i(K)| + \rho \cdot \max_{i=1}^{n} \max_{k=1}^{m} |X_0(K) - X_i(K)|}
\]

Which (\(k=1,2,... m\))

In the above formula, \(\rho\) is the resolution coefficient, which is taken within (0,1). If \(\rho\) is smaller, the difference between the correlation coefficients is greater, and the discrimination ability is stronger. The value is 0.5 under normal research conditions

(6) Calculate the value of the correlation coefficient. For each evaluation object, calculate the mean value of the correlation coefficient of its \(m\) indicators and the corresponding elements of the reference sequence, to reflect the correlation between each evaluation object and the reference sequence, and describe it as the correlation sequence, which is recorded as (2):

\[
r_w = \frac{1}{m} \sum_{k=1}^{m} W_k \cdot \zeta(k)
\]

(Which \(k=1,2,... m\))

If each indicator plays a different role in the comprehensive evaluation, the average value of the correlation coefficient can be calculated

(7) According to the correlation order of each observation object, comprehensive evaluation results are obtained.

4 RESULTS

The Informatization evaluation index system of this paper will adopt the research evaluation system [12] and combine it with the actual situation of SMEs in this survey, with minor modifications, as shown in Table 1:
Table 1. The level of the IRM’s Evaluation Index.

<table>
<thead>
<tr>
<th>Character</th>
<th>Index</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td>The proportion of the investment in IRM</td>
<td>Reflects the informationize investment</td>
</tr>
<tr>
<td>X2</td>
<td>Every one hundred people computer ownership</td>
<td>Reflects the proportion of the hardware facilities</td>
</tr>
<tr>
<td>X3</td>
<td>The rate of the connecting Network</td>
<td>Reflects the rate of connecting Network</td>
</tr>
<tr>
<td>X4</td>
<td>The rate of the information coverage</td>
<td>Reflects the application status of the market, sales, and technology in IRM</td>
</tr>
<tr>
<td>X5</td>
<td>The proportion of the IRM talents</td>
<td>Reflects the proportion of talents in IRM</td>
</tr>
<tr>
<td>X6</td>
<td>The proportion of IRM security spending</td>
<td>Reflect the proportion of the security spending in whole IRM spending</td>
</tr>
</tbody>
</table>

(1) To establish a six matrix by the above index:

Table 2. Sixth order matrix.

<table>
<thead>
<tr>
<th>Number (%)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td>8</td>
<td>9</td>
<td>6</td>
<td>7</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>X2</td>
<td>25</td>
<td>67</td>
<td>33</td>
<td>44</td>
<td>61</td>
<td>35</td>
</tr>
<tr>
<td>X3</td>
<td>38</td>
<td>84</td>
<td>57</td>
<td>54</td>
<td>78</td>
<td>66</td>
</tr>
<tr>
<td>X4</td>
<td>23</td>
<td>68</td>
<td>34</td>
<td>45</td>
<td>72</td>
<td>32</td>
</tr>
<tr>
<td>X5</td>
<td>13</td>
<td>53</td>
<td>26</td>
<td>67</td>
<td>54</td>
<td>37</td>
</tr>
<tr>
<td>X6</td>
<td>7</td>
<td>21</td>
<td>11</td>
<td>26</td>
<td>27</td>
<td>10</td>
</tr>
</tbody>
</table>

(2) Determines the reference sequence and original data standardization by the maximum

\{X_0\} = \{10, 67, 84, 72, 67, 27\}

Table 3. Data Standardization.

<table>
<thead>
<tr>
<th>Number (%)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td>0.8</td>
<td>0.9</td>
<td>0.6</td>
<td>0.7</td>
<td>1</td>
<td>0.6</td>
</tr>
<tr>
<td>X2</td>
<td>0.37</td>
<td>1</td>
<td>0.5</td>
<td>0.66</td>
<td>0.91</td>
<td>0.52</td>
</tr>
<tr>
<td>X3</td>
<td>0.45</td>
<td>1</td>
<td>0.68</td>
<td>0.64</td>
<td>0.92</td>
<td>0.79</td>
</tr>
<tr>
<td>X4</td>
<td>0.32</td>
<td>0.94</td>
<td>0.47</td>
<td>0.63</td>
<td>1</td>
<td>0.44</td>
</tr>
<tr>
<td>X5</td>
<td>0.19</td>
<td>0.79</td>
<td>0.39</td>
<td>1</td>
<td>0.8</td>
<td>0.55</td>
</tr>
<tr>
<td>X6</td>
<td>0.26</td>
<td>0.78</td>
<td>0.41</td>
<td>0.96</td>
<td>1</td>
<td>0.37</td>
</tr>
</tbody>
</table>
(3) Calculates \( |X_0(K) - X(K)| \) absolute difference value and establishes the matrix:

**Table 4.** Matrix established by Absolute Value Difference.

<table>
<thead>
<tr>
<th>Number (%)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td>0.2</td>
<td>0.1</td>
<td>0.4</td>
<td>0.3</td>
<td>0</td>
<td>0.4</td>
</tr>
<tr>
<td>X2</td>
<td>0.63</td>
<td>0</td>
<td>0.6</td>
<td>0.34</td>
<td>0.09</td>
<td>0.48</td>
</tr>
<tr>
<td>X3</td>
<td>0.55</td>
<td>0</td>
<td>0.32</td>
<td>0.36</td>
<td>0.07</td>
<td>0.21</td>
</tr>
<tr>
<td>X4</td>
<td>0.68</td>
<td>0.06</td>
<td>0.53</td>
<td>0.38</td>
<td>0</td>
<td>0.56</td>
</tr>
<tr>
<td>X5</td>
<td>0.81</td>
<td>0.21</td>
<td>0.61</td>
<td>0</td>
<td>0.2</td>
<td>0.45</td>
</tr>
<tr>
<td>X6</td>
<td>0.74</td>
<td>0.22</td>
<td>0.59</td>
<td>0.04</td>
<td>0</td>
<td>0.63</td>
</tr>
</tbody>
</table>

(4) Determines the \( \min_{i=1}^{n} \min_{k=1}^{m} |X_0(K) - X_i(K)| \) and \( \max_{i=1}^{n} \max_{k=1}^{m} |X_0(K) - X_i(K)| \)

\[
\min_{i=1}^{n} \min_{k=1}^{m} |X_0(K) - X_i(K)| = \min_{i=1}^{n} (0.20, 0, 0.32, 0, 0, 0.21) = 0
\]

\[
\max_{i=1}^{n} \max_{k=1}^{m} |X_0(K) - X_i(K)| = \max_{k=1}^{m} (0.81, 0.22, 0.61, 0.38, 0.20, 0.63) = 0.81
\]

(5) Calculating correlation coefficient

\[
\rho = 0.5 \cdot \xi_1 = (0 + 0.5 \cdot 0.81) / (0.2 + 0.5 \cdot 0.81) \approx 0.669
\]

Similarly all the other value can be calculated and builds the matrix of Calculating correlation coefficient:

**Table 5.** Resulting Correlation Coefficient.

<table>
<thead>
<tr>
<th>Number (%)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td>0.669</td>
<td>0.802</td>
<td>0.503</td>
<td>0.570</td>
<td>1</td>
<td>0.503</td>
</tr>
<tr>
<td>X2</td>
<td>0.393</td>
<td>1</td>
<td>0.403</td>
<td>0.534</td>
<td>0.818</td>
<td>0.458</td>
</tr>
<tr>
<td>X3</td>
<td>0.424</td>
<td>1</td>
<td>0.559</td>
<td>0.536</td>
<td>0.853</td>
<td>0.659</td>
</tr>
<tr>
<td>X4</td>
<td>0.373</td>
<td>0.871</td>
<td>0.433</td>
<td>0.526</td>
<td>1</td>
<td>0.420</td>
</tr>
<tr>
<td>X5</td>
<td>0.333</td>
<td>0.659</td>
<td>0.399</td>
<td>1</td>
<td>0.669</td>
<td>0.458</td>
</tr>
<tr>
<td>X6</td>
<td>0.353</td>
<td>0.648</td>
<td>0.407</td>
<td>0.910</td>
<td>1</td>
<td>0.391</td>
</tr>
</tbody>
</table>
5 DISCUSSION

Calculates the average of the correlation coefficient
\[ r_{01} = (0.669 + 0.393 + 0.424 + 0.373 + 0.333 + 0.353) / 6 = 0.424 \]
\[ r_{02} = (0.802 + 1 + 1 + 0.871 + 0.659 + 0.648) / 6 = 0.830 \]

The same procedure may be easily reached:
\[ r_{03} = 0.451 \quad r_{04} = 0.679 \quad r_{05} = 0.890 \quad r_{06} = 0.482 \]

To sum up, the ranking of information resource management level of these six enterprises is (5) E Electronics Industry Co., Ltd., (2) Guangzhou B Technology Co., Ltd., (4) Guangzhou D Plastic Technology Co., Ltd., (6) F Shoes Co., Ltd., (3) Guangzhou C Industry Co., Ltd., and (1) Guangzhou A Solid Wood Furniture Co., Ltd. The sorting result is consistent with the enterprise profit growth rate fed back by the enterprise.

The above empirical research shows that there is a positive relationship between the level of information capital management and the profit growth rate of SMEs. Nowadays, many small and medium-sized enterprises are faced with problems such as a lack of funds and resources within the enterprise, but the author believes that the investment of small and medium-sized enterprises in information resource management can be said to be crucial because these investments can achieve profit synchronous growth. Therefore, SMEs should strengthen their confidence and improve their information resource management level.

6 CONCLUSION

In the future, small and medium-sized enterprises can continuously improve their information management level with big data as support. Promote the construction and development of small and medium-sized enterprises through the establishment of the information management model system. While improving information management, do a good job in the construction of the basic platform of information management to ensure the improvement of the information level of small and medium-sized enterprises. According to the typical Nolan six-stage model (Nolan) of enterprise informatization development stage research, enterprise informatization can be divided into six stages: initial stage, expansion stage, control stage, unified stage, data management stage, and mature stage. The first three stages have the characteristics of the computer age, and the last three stages have the characteristics of the information age. The turning point is the time for information resource planning. Nolan stressed that any organization must develop from one stage to the next when implementing computer-based information systems, and cannot achieve leapfrog development. In the future, SMEs can improve the informatization level from the following aspects.

(1) Optimize the construction of enterprise informatization basic platform. Under the background of the big data era, SMEs should establish sound data management thinking, focusing on the construction of high-quality information basic management platforms. At present, the information infrastructure of small and medium-sized enterprises often cannot fully meet the
requirements of the big data era. Enterprises should pay attention to the construction of information talents and technical equipment to provide support for the construction of information management infrastructure platforms. It is necessary to improve the data processing, storage, and analysis of information systems, effectively integrate and upgrade resources, so that enterprises can obtain support in management decisions, promote further development of enterprises, and obtain corresponding power support.

(2) Promote the standardization and integration of enterprise information systems to achieve the stable construction of enterprise information management systems, it is necessary to design from multiple aspects, such as top-level design, process management control, and terminal, to achieve data exchange and association at different levels, form a system integration model, and make enterprises obtain space for improvement in the development process.

(3) Create a perfect information data network system, and make a good investment in informatization. The enterprise management should take the product as the center, and the management mode should take the service as the core. Fundamentally, the new management model should start from the following aspects: form a data information network model, which mainly covers the product information in the production process of enterprises, and infiltrate employees and service content into it. Employees should be assessed to strengthen their comprehensive ability so that they can establish correct concepts and cognition, form a new mode of thinking, give play to the actual value of data, and provide good support for enterprise information construction through sorting out and screening information resources. Under the condition of fully implementing the utilization of resources, we should attach importance to the personalized characteristics of information networks within enterprises and strengthen the updating and implementation of enterprise management models. In addition, we should do a good job of investing in information technology. Small and medium-sized enterprises should form a sound financing plan according to their specific conditions, invest information funds in an all-around way, highlight vitality and vitality, and better face market competition.

(4) Strengthen the training and construction of enterprise informatization talents. For enterprises, the acquisition of informatization talents is a difficult point as well as a key point. First of all, enterprises should, through close cooperation with colleges and universities, and the implementation of school-enterprise cooperation, use the excellent talents of the school as the support to tap and cultivate the information talents needed for their development. Secondly, based on the current development situation, enterprises should look for information management talents that can be trained from enterprises to improve the information management level and ability to existing managers. Thirdly, enterprises should be good at using external excellent information management resources. For example, some enterprises that do not have the strength to build an information management team suitable for their development can temporarily hire some external information consulting and management companies to use their professional knowledge and management systems to improve their management effects.

Acknowledgements: The authors would like to express thankfulness and gratitude to Guangzhou City University of Technology, for providing adequate support during this research work. Most importantly, I wish to thank my loving and supportive family and students, who provide unending inspiration to complete this research.
Funding: This research is supported by the Guangdong Provincial University Online Course Committee Fund, China, Number: 2022ZXKC589, and 2022 Research Project of the Professional Committee of Private Higher Education of Guangdong Higher Education Association, China, Number: 2022MBGJ044.

REFERENCES

Asset Pricing in China’s Stock Market

Lu Yu

yulu2023@yahoo.com

University of California Davis, 1 shields Ave, Davis, 95618, the United States

Abstract: This paper tests four main-stream asset pricing model with China’s stock market information. We use stock return data from CSMAR and constructed China Stock Market Factor. The result suggests that the Fama and French (2014) five-five model is a better description of Chinese market than the Fama and French (1993) three-factor model.

Keywords: Asset pricing model, Fama and French five-factor model, Chinese stock market.

1 INTRODUCTION

With the development of society, people gradually realize the importance of data. Investors try to find factors related to stock return with several datasets. China has already been the second greatest stock market around the world. The passage attempts to figure out which asset model is the best description of Chinese market. We examine four main-stream asset pricing model. The CAMP asset pricing model, the Fama and French (1993) three-factor model, the Carhart (1997) four-factor model, and the Fama and French (2014) five-five model. These inclusive five factors are Market risk factor (RM-RF), size risk factor (SMB) and Book-to-Market ratio factor (HML) Profitability Factor (RMW), and Investment Pattern Factor (CMA). We use data from CSMAR and sub-database, and the other one is China Stock Market Factor Database. We have discovered that the Fama and French (2014) five-factor model provides a more accurate depiction of the Chinese market compared to the Fama and French (1993) three-factor model.

2 LITERATURE REVIEW

There has been extensive studies in the asset pricing literatures, for example the renowned Capital Asset Pricing Model, the CAMP asset pricing model, the Fama and French (1993) three-factor model, the Carhart (1993) four-factor model, and the Fama and French (2015) five-five model. Fama and French (2015) focus on a five-factor asset pricing model, capturing the size, value, profitability, and investment patterns in average stock returns. This five-factor model performs better than the traditional Fama-French three-factor model (Fama and French, 1993). However, the five-factor model fails to capture the low average returns on small stocks whose returns behave like those of firms that invest a lot despite low profitability. The authors suggest that HML is a redundant factor in the sense that its high average return is fully captured by its exposures to Rm-Rf, SMB, and especially RMW and CMA. Also, the tests indicate that a four-factor model dropping HML performs as well as the five-factor model. [7] (Mosoeu and Kodongo, BDEDM 2023, January 06-08, Changsha, People’s Republic of China Copyright © 2023 EAI DOI 10.4108/eai.6-1-2023.2330345)
2020) focus on the Fama-French five-factor model and emerging market equity returns. They test the model on average stock returns for selected emerging and developed equity market. They find that the profitability factor is the most useful for explaining the cross-section emerging markets equity returns. Another important result is that the average returns of stocks of large-size firms appear to exceed those of stocks of small-size firms and that returns on stocks of growth firms exceed returns on stocks of value firms.

Nevertheless, researchers have been consistently exploring additional factors that would contribute to explain asset prices, and also across different locations. [4] (Guo, Zhang, Zhang and Zhang, 2017) focus on the five-factor asset pricing model tests for the Chinese stock market. The test results suggest that strong size, value, and profitability patterns in average returns, but weak investment pattern. They find that the profitability factor significantly improves the description of average return. The test results are consistent with Chen et al. (2010), who show that many anomaly variables which are efficient in the U.S. market but do not affect the average returns of the Chinese market, except the obvious value effect. [9] (Liu, Stambaugh and Yuan, 2019) focus on size and value factors in China. They find that the three-factor model based on Chinese stock market, CH-3, dominates the traditional FF-3 model, where in the latter is based on the BM value factor. The CH-3 model prices both the size and value factors in FF-3. The three-factor model strongly dominates a model formed by just replicating the Fama and French (1993) procedure in China and explains most reported anomalies in the Chinese stock market, including profitability and volatility. [8] (Li and Rao, 2022) focus on a revised model, accounting for unique features of Chinese market, and evaluate the performance of competing asset price models. Li and Rao provide an effective benchmark model for empirical asset pricing in the Chinese stock market, which demonstrates that the propensity of firms to engage in reverse merges has sharply decreased in recent years. [1] (Chen, Glabadanidis and Sun 2022) mainly focus on the five-factor asset pricing, short-term reversal, and ownership structure. They find that the five-factor asset pricing model proposed by [2] (Fama and French 2015) is a better description of the Chinese stock market return than the three-factor asset pricing model. Moreover, they propose a short-term-reversal (STR) is highly significant, which substantially improves the pricing ability of three- and five- asset pricing models in explaining popular stock portfolio returns as well as Chinese mutual funds’ returns. [10] (Fama and French 2017) employ the international test of a five-factor asset pricing model. Average stock returns for North America, Europe, and Asia Pacific increase with the book-to-market ratio (B/M) and profitability and are negatively related to investment. In the case of Japan, there is a significant connection between average returns and book-to-market ratio (B/M), while average returns display minimal correlation with profitability or investment. By incorporating factors of profitability and investment into the Fama and French (1993) three-factor model, a five-factor model successfully encompasses the observed patterns in average returns.

3 DATA AND VARIABLES

There are two main sources of data for this survey. One is CSMAR and sub-database, and the other one is China Stock Market Factor Database. We focus on the A-share broad market factors and select the 2*2 portfolio division method for market risk factor construction. Ultimately, the full data contains five-year sample from 12 October 2017 to 11 October 2022. It contains 1214-day observation and 5149 firms. The overall data set contains 4901762 firm-day observations.
In the later robustness examination, we also use the weekly observation. The week sample contains 253 weeks and 5145 firms, in total of 2878834 firm-week observations.

The market risk premium is that we compute the disparity between the daily market return and the daily risk-free rate, accounting for the reinvestment of cash dividends. Specifically, we evaluate market risk using two approaches: the trading value weighted average of shares listed in the market and the market capitalization value weighted average. We refer to the former risk premium measure as mk\_rf1 and the latter as mk\_rf2. The risk-free rate is based on the benchmark deposits published by the Central Bank of China in March.

The size factor, alternatively known as the market capitalization factor is calculated the difference between the return of a small-cap portfolio and a large-cap portfolio. Again, we denote the trading value-weighted size factor as SMB1, and the market capitalization weighting measure as SMB2. Similar fashion also applies to the following measures.

The book-to-market factor, which calculates the difference between the return on a portfolio with a high book-to-market ratio and a portfolio with a low book-to-market ratio. Particularly, we measure HML in two ways, one is the liquidity weighted and the other one is the total market value weighted. We name the former HML measure as HML1 and the latter one as HML2. Therefore, the portfolio investment return for the former one is using market capitalization weighting, and the portfolio investment return for the latter one is the total market value weighting.

The profitability factor is calculating the difference between the return on a high-profit portfolio of stocks and a low-profit portfolio. Specifically, there are two types of RMW. RMW1 represents the market capitalization weighted and RMW2 is total market capitalization weighted. Portfolio investment returns in the former one is calculated using market capitalization weighting., and for the latter one is using total market capitalization weighting.

The investment pattern factor is calculated as the difference between the return of a low and high investment ratio stock portfolio. Concretely, we measure the investment pattern factor in two ways, one is the market capitalization weighted average, and the other one is total market capitalization weighted average of shares listed in the market. We denote the former one as CMA1 and the latter one as CMA2.

Dretwd denotes the daily stock return with considering of reinvestment of cash dividends. It is measured as the stock offer price minus the previous closing price plus the reinvestment return from dividends by earning a market risk-free return. On the other hand, Dretnd captures the daily stock return without considering of reinvestment of cash dividends.

Wretwd represents the weekly individual stock return with considering of cash dividend reinvestment. It is measured as the stock offer price minus the previous closing price plus the reinvestment return from dividends by earning a market risk-free return. On the other hand, wretnd indicates weekly individual stock return without considering of cash dividend reinvestment.

4 METHODOLOGY

In this paper, we examine four main-stream asset pricing model. The CAMP asset pricing model,
the Fama and French (1993) three-factor model, the Carhart (1997) four-factor model, and the Fama and French (2014) five-factor model. These inclusive five factors are Market risk factor (RM-RF), size risk factor (SMB) and Book-to-Market ratio factor (HML) Profitability Factor (RMW), and Investment Pattern Factor (CMA). Furthermore, equations (1) to (4) list the standard asset pricing regression models.

\[ R_{it} - R_{ft} = \alpha_i + \beta_1(R_{mt} - R_{ft}) + \epsilon_{it} \]  
\[ R_{it} = \alpha_i + \beta_1(R_{mt} - R_{ft}) + \beta_2 SMB_t + \beta_3 HML_t + \epsilon_{it} \]  
\[ R_{it} = \alpha_i + \beta_1(R_{mt} - R_{ft}) + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 RMW_t + \epsilon_{it} \]  
\[ R_{it} = \alpha_i + \beta_1(R_{mt} - R_{ft}) + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 RMW_t + \beta_5 CMA_t + \epsilon_{it} \]

\( R_{it} \) captures the total return of a stock \( i \) at time \( t \), and \( R_{ft} \) is risk free rate of return at time \( t \). Moreover, \( R_{mt} \) represents total market portfolio return at time \( t \). \( R_{it} - R_{ft} \) means the firm’s expected excess return. Additionally, \( R_{mt} - R_{ft} \) is excess return on the market portfolio (index); SMB\( _t \) stands for size premium (small minus big); HML\( _t \) is the representation of value premium (high minus low). In addition, RMW captures the returns on diversified portfolios of stock with robust (high and steady) minus weak (low) profitability, hence represents the momentum factor. Finally, CMA captures returns on diversified portfolios of the stocks of low reinvestment ratio and high reinvestment ratio investment firms, which represents conservative and aggressive dividend-payout ratios of the companies. \( \beta_1 \) to \( \beta_5 \) are the factor coefficients that we want to estimate based on the Chinese market data.

5 EMPIRICAL RESULTS

Table 1 presents the summary statistics of variables that used in the study, including the number of observation (N), sample average (Mean), standard deviation (SD), minimum (Min), median (p50), and maximum (max) values of variables. As shown in the table, the total number of daily stock return, dretwd, is 4901762, with the mean of .0004402, standard deviation of .0434482, min of -0.895775, median of 0 and max of 19.42581. The standard deviation is much higher than the mean, which is more than ten times of the mean. In other words, the portfolio fluctuates a lot. For the dretnd, the total number is the same as dretwd, which is 4901762. The mean is 0.000396 and the standard deviation is 0.0434689. Both min, p50 and max are the same as the former variable. When it comes to mk_rf1, the total number of observations is 4901762 and the mean is 0.0001183. The standard deviation is 0.117876 and the min is -0.079134. P50 is 0.000586, which is slightly higher than the former one, but max is 0.056292 that is much smaller. Also, the portfolio fluctuates a lot because the standard deviation is higher than the mean. In terms of hml1, the mean is -8.23e-06 and the standard deviation is .0052621, both of which are relatively lower than the previous variables. Min is -0.19516 and p50 is -0.000317. Max is 0.17008. The portfolio shares similar moves with all the previous variables. For cma1,
the mean is -.0000315 and the standard deviation is .0033758. The min, p50 and max are 
-.012392, -.000084, and .013065.

Table 1. Summary Statistics: Daily Returns

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>p50</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>dretwd</td>
<td>4901762</td>
<td>.0004402</td>
<td>.0434482</td>
<td>-.895775</td>
<td>0</td>
<td>19.42581</td>
</tr>
<tr>
<td>dretnd</td>
<td>4901762</td>
<td>.000396</td>
<td>.0434689</td>
<td>-.895775</td>
<td>0</td>
<td>19.42581</td>
</tr>
<tr>
<td>mk_rf1</td>
<td>4901762</td>
<td>.001183</td>
<td>.0117876</td>
<td>-.079134</td>
<td>.000586</td>
<td>.056292</td>
</tr>
<tr>
<td>smb1</td>
<td>4901762</td>
<td>.000583</td>
<td>.0081276</td>
<td>-.035087</td>
<td>.000573</td>
<td>.027585</td>
</tr>
<tr>
<td>hml1</td>
<td>4901762</td>
<td>-8.23e-06</td>
<td>.0052621</td>
<td>-.019516</td>
<td>-.000317</td>
<td>.017008</td>
</tr>
<tr>
<td>rmw1</td>
<td>4901762</td>
<td>.0001432</td>
<td>.0038559</td>
<td>-.016499</td>
<td>.000173</td>
<td>.016747</td>
</tr>
<tr>
<td>cma1</td>
<td>4901762</td>
<td>-.000315</td>
<td>.0033758</td>
<td>-.012392</td>
<td>-.000848</td>
<td>.13065</td>
</tr>
<tr>
<td>mk_rf2</td>
<td>4901762</td>
<td>.002838</td>
<td>.01214</td>
<td>-.08082</td>
<td>.008851</td>
<td>.056508</td>
</tr>
<tr>
<td>smb2</td>
<td>4901762</td>
<td>.000197</td>
<td>.0076889</td>
<td>-.030535</td>
<td>.000519</td>
<td>.025245</td>
</tr>
<tr>
<td>hml2</td>
<td>4901762</td>
<td>-.000077</td>
<td>.0049466</td>
<td>-.018893</td>
<td>.000376</td>
<td>.016418</td>
</tr>
<tr>
<td>rmw2</td>
<td>4901762</td>
<td>.001298</td>
<td>.0040289</td>
<td>-.017676</td>
<td>.00176</td>
<td>.017886</td>
</tr>
<tr>
<td>cma2</td>
<td>4901762</td>
<td>-.000158</td>
<td>.0034831</td>
<td>-.01306</td>
<td>-.000911</td>
<td>.13941</td>
</tr>
</tbody>
</table>

This table presents the summary statistics of variables that used in the study. It includes the number of observation (N), sample average (Mean), standard deviation (SD), minimum (Min), median (p50), and maximum (max) values of variables.

This table presents the summary statistics of variables that used in the study. It includes the number of observation (N), sample average (Mean), standard deviation (SD), minimum (Min), median (p50), and maximum (max) values of variables. The total number of observations is 1002357. For the wretwd, the mean is 0.000116. Moreover, the standard deviation is 0.0679955. Min, p50m, and max are -.895775, -.002967, and 3.872825. The standard of deviation is much higher than the mean, so the portfolio fluctuates a lot. When it comes to wretnd, the mean is .0001423 and the standard deviation is .0680482. It is obvious that the standard deviation is still higher than the mean, so the moves of the portfolio of wretnd is similar with wretwd. In terms of mk_rf1, The mean is .0001177 and the standard deviation is .0244852. The standard deviation is still higher than the mean, so the portfolio of mk_rf1 fluctuates a lot. The min, p50, and max are -.099403, .000884, and .082596. Concerning to hml1, the mean and standard deviation are -.000103 and .0113316. Apparently, the portfolio of hml1 fluctuates a lot. The min, p50, and max are -.028418, -.001279, and .037644.

Table 2. Summary Statistics: Weekly Returns

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>p50</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>wretwd</td>
<td>1002357</td>
<td>.000116</td>
<td>.0679955</td>
<td>-.895775</td>
<td>-.002967</td>
<td>3.872825</td>
</tr>
<tr>
<td>wretnd</td>
<td>1002357</td>
<td>.0001432</td>
<td>.0680482</td>
<td>-.895775</td>
<td>-.002964</td>
<td>3.872825</td>
</tr>
<tr>
<td>mk_rf1</td>
<td>1003792</td>
<td>.0001177</td>
<td>.0244852</td>
<td>-.099403</td>
<td>.000602</td>
<td>.082596</td>
</tr>
<tr>
<td>smb1</td>
<td>1003792</td>
<td>.0007365</td>
<td>.0180571</td>
<td>-.067276</td>
<td>.000904</td>
<td>.054241</td>
</tr>
</tbody>
</table>
This table presents the summary statistics of variables that used in the study. It includes the number of observation (N), sample average (Mean), standard deviation (SD), minimum (Min), median (p50), and maximum (max) values of variables.

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Median</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>hml1</td>
<td>1003792</td>
<td>-0.00103</td>
<td>0.0113316</td>
<td>-0.028418</td>
<td>-0.001279</td>
<td>0.037644</td>
</tr>
<tr>
<td>rmw1</td>
<td>1003792</td>
<td>0.0004411</td>
<td>0.0091104</td>
<td>-0.033482</td>
<td>0.000708</td>
<td>0.032173</td>
</tr>
<tr>
<td>cma1</td>
<td>1003792</td>
<td>-0.000097</td>
<td>0.0073286</td>
<td>-0.023371</td>
<td>-0.000723</td>
<td>0.022254</td>
</tr>
<tr>
<td>mk_rf2</td>
<td>1003792</td>
<td>-0.0000388</td>
<td>0.0250088</td>
<td>-0.096956</td>
<td>0.000764</td>
<td>0.087175</td>
</tr>
<tr>
<td>smb2</td>
<td>1003792</td>
<td>0.000514</td>
<td>0.0170897</td>
<td>-0.028598</td>
<td>-0.001666</td>
<td>0.040238</td>
</tr>
<tr>
<td>hml2</td>
<td>1003792</td>
<td>-0.0004493</td>
<td>0.0094793</td>
<td>-0.034569</td>
<td>0.000976</td>
<td>0.031089</td>
</tr>
<tr>
<td>rmw2</td>
<td>1003792</td>
<td>0.0003338</td>
<td>0.0075386</td>
<td>-0.022417</td>
<td>-0.000567</td>
<td>0.023441</td>
</tr>
<tr>
<td>cma2</td>
<td>1003792</td>
<td>-8.83e-06</td>
<td>0.0075386</td>
<td>-0.022417</td>
<td>-0.000567</td>
<td>0.023441</td>
</tr>
</tbody>
</table>

This table presents the summary statistics of variables that used in the study. It includes the number of observation (N), sample average (Mean), standard deviation (SD), minimum (Min), median (p50), and maximum (max) values of variables.

**Table 3. Regression Results: Trading-Valued Weighted Return with Dividends**

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>dretwd</td>
<td>dretwd</td>
<td>dretwd</td>
<td>dretwd</td>
<td>wretwd</td>
<td>wretwd</td>
<td>wretwd</td>
<td>wretwd</td>
</tr>
<tr>
<td>mk_rf1</td>
<td>1.098***</td>
<td>1.032***</td>
<td>1.028***</td>
<td>1.028***</td>
<td>0.565***</td>
<td>0.516***</td>
<td>0.514***</td>
<td>0.516***</td>
</tr>
<tr>
<td></td>
<td>(1027.68)</td>
<td>(912.86)</td>
<td>(907.07)</td>
<td>(906.00)</td>
<td>(224.11)</td>
<td>(194.81)</td>
<td>(190.28)</td>
<td>(190.95)</td>
</tr>
<tr>
<td>smb1</td>
<td>0.836***</td>
<td>0.759***</td>
<td>0.757***</td>
<td>0.517***</td>
<td>0.504***</td>
<td>0.496***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(564.09)</td>
<td>(342.99)</td>
<td>(340.77)</td>
<td>(151.95)</td>
<td>(97.85)</td>
<td>(96.24)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hml1</td>
<td>-0.0706***</td>
<td>-0.125***</td>
<td>-0.131***</td>
<td>-0.119***</td>
<td>-0.126***</td>
<td>-0.210***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-28.76)</td>
<td>(-45.95)</td>
<td>(-44.57)</td>
<td>(-20.92)</td>
<td>(-20.70)</td>
<td>(-30.70)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rmw1</td>
<td>-0.222***</td>
<td>-0.211***</td>
<td>-0.133***</td>
<td>-0.119***</td>
<td>-0.126***</td>
<td>-0.210***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-47.02)</td>
<td>(-41.29)</td>
<td>(-3.32)</td>
<td>(7.92)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cma1</td>
<td>0.0270***</td>
<td>0.305***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(5.57)</td>
<td>(26.84)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Const</td>
<td>-0.0000781</td>
<td>-0.000161***</td>
<td>-0.000132**</td>
<td>-0.000133**</td>
<td>-0.000922***</td>
<td>-0.00136***</td>
<td>-0.00133***</td>
<td>-0.00137***</td>
</tr>
<tr>
<td></td>
<td>(-1.34)</td>
<td>(-2.95)</td>
<td>(-2.36)</td>
<td>(-2.41)</td>
<td>(-7.69)</td>
<td>(-11.43)</td>
<td>(-11.26)</td>
<td>(-11.62)</td>
</tr>
<tr>
<td>N</td>
<td>4901762</td>
<td>4901762</td>
<td>4901762</td>
<td>4901762</td>
<td>1002357</td>
<td>1002357</td>
<td>1002357</td>
<td>1002357</td>
</tr>
<tr>
<td>R2</td>
<td>0.177</td>
<td>0.229</td>
<td>0.230</td>
<td>0.230</td>
<td>0.0477</td>
<td>0.0718</td>
<td>0.0718</td>
<td>0.0724</td>
</tr>
</tbody>
</table>

In Table 3, the trading-valued returns including dividends are presented. The left panel consists of four columns reporting results based on daily returns, while the right panel consists of four columns reporting results based on weekly returns. From column 1 to column 4 in the left panel (corresponding to columns 5 to 8 in the right panel), a gradual regression analysis is conducted, where stock returns are regressed on the market-risk premium, size factor, book-to-market factor, profitability factor, and investment pattern factor.
First, all factors illustrate significant explanatory power in explaining stock returns, with the significance level of 1%. Second, for each of the series of the mk_rf1, smb1, and hml1, for example, the trading-valued weighted return with dividends of mkt_rf1 decreases from 1.098 to 0.516. The smb1 decreases from 0.836 to 0.496. The hml1 decreases from -0.0706 to -0.210. The sign in front of mkt_rf1, smb1, cma1 are all positive, which means with these market factors increase, stock returns also increase. In addition, it represents the smaller the company is, the higher the return is. The investment pattern factor gets growing, the return will get higher. However, for each of the series of the cma1 and rmw1, the effect on trading-valued weighted return with dividends from cma1 increases from 0.0270 to 0.305. Similarly, the effect of rmw1 increases from -0.222 to 0.0879. The sign in front of hml1 and rmw1 are negative. In other words, as book-to-market factor be higher, the lower returns it is. Also, the high of profitability factor, the lower return it is.

These overall findings are consistent with the five-factor asset pricing model, short-term reversal, and ownership structure – the case of China (Chen, Glabadanidis and Sun, 2022). In other words, the five-factor asset pricing model can be explained by Chinese asset returns.

Table 4. Regression Results: Trading-Valued Return without Dividend

<table>
<thead>
<tr>
<th></th>
<th>detdnd</th>
<th>detdnd</th>
<th>detdnd</th>
<th>detdnd</th>
<th>wretnd</th>
<th>wretnd</th>
<th>wretnd</th>
<th>wretnd</th>
</tr>
</thead>
<tbody>
<tr>
<td>mk_rf1</td>
<td>1.098***</td>
<td>1.032***</td>
<td>1.028***</td>
<td>1.028***</td>
<td>0.565***</td>
<td>0.516***</td>
<td>0.514***</td>
<td>0.516***</td>
</tr>
<tr>
<td></td>
<td>(1026.19)</td>
<td>(911.64)</td>
<td>(905.86)</td>
<td>(904.77)</td>
<td>(223.98)</td>
<td>(194.68)</td>
<td>(190.16)</td>
<td>(190.83)</td>
</tr>
<tr>
<td>smb1</td>
<td>0.836***</td>
<td>0.759***</td>
<td>0.758***</td>
<td>0.517***</td>
<td>0.504***</td>
<td>0.497***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(563.57)</td>
<td>(342.68)</td>
<td>(340.49)</td>
<td>(151.88)</td>
<td>(97.82)</td>
<td>(96.21)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hml1</td>
<td>-0.0695***</td>
<td>-0.123***</td>
<td>-0.129***</td>
<td>-0.119***</td>
<td>-0.126***</td>
<td>-0.210***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-28.24)</td>
<td>(-45.46)</td>
<td>(-44.02)</td>
<td>(-20.93)</td>
<td>(-20.71)</td>
<td>(-30.70)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rmw1</td>
<td>-0.222***</td>
<td>-0.211***</td>
<td>-0.211***</td>
<td>-0.0335***</td>
<td>0.0881***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-46.96)</td>
<td>(-41.33)</td>
<td>(-3.31)</td>
<td>(7.93)</td>
<td>(7.93)</td>
<td>(7.93)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cma1</td>
<td>0.0258***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(5.31)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_cons</td>
<td>-0.000125**</td>
<td>-0.000210***</td>
<td>-0.000177***</td>
<td>-0.000178***</td>
<td>-0.000090***</td>
<td>-0.000133***</td>
<td>-0.000131***</td>
<td>-0.000135** **</td>
</tr>
<tr>
<td></td>
<td>(-2.10)</td>
<td>(-3.75)</td>
<td>(-3.16)</td>
<td>(-3.23)</td>
<td>(-7.51)</td>
<td>(-11.25)</td>
<td>(-11.08)</td>
<td>(-11.44)</td>
</tr>
<tr>
<td>N</td>
<td>4901762</td>
<td>4901762</td>
<td>4901762</td>
<td>4901762</td>
<td>1002357</td>
<td>1002357</td>
<td>1002357</td>
<td>1002357</td>
</tr>
<tr>
<td>r2_o</td>
<td>0.177</td>
<td>0.229</td>
<td>0.229</td>
<td>0.229</td>
<td>0.0476</td>
<td>0.0717</td>
<td>0.0717</td>
<td>0.0724</td>
</tr>
</tbody>
</table>

In Table 4, the trading-weighted returns excluding dividends are presented for the market-risk premium, size factor, book-to-market factor, profitability factor, and investment pattern factor. Firstly, it is noteworthy that all factors exhibit significant explanatory power for returns at a 1% significance level, thus confirming the robustness of the five-factor pricing model. Secondly, when analyzing mk_rf1, smb1, and hml1, a declining return pattern is observed from (1) to (8). Conversely, for cma1 and rmw1, an increasing trend in returns is observed from (1) to (8).
Table 5. Regression Results: Capitalization Weighted Return with Dividend

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>mk_rf2</td>
<td>1.085***</td>
<td>1.019***</td>
<td>1.014***</td>
<td>1.015***</td>
<td>0.577***</td>
<td>0.526***</td>
<td>0.522***</td>
<td>0.526***</td>
</tr>
<tr>
<td></td>
<td>(1057.48)</td>
<td>(915.11)</td>
<td>(908.51)</td>
<td>(907.88)</td>
<td>(233.52)</td>
<td>(197.93)</td>
<td>(192.50)</td>
<td>(193.82)</td>
</tr>
<tr>
<td>smb2</td>
<td>0.829***</td>
<td>0.732***</td>
<td>0.730***</td>
<td>0.509***</td>
<td>0.481***</td>
<td>0.475***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(522.35)</td>
<td>(310.13)</td>
<td>(327.46)</td>
<td>(327.46)</td>
<td>(139.82)</td>
<td>(94.66)</td>
<td>(93.44)</td>
<td></td>
</tr>
<tr>
<td>hml2</td>
<td>0.0769***</td>
<td>-0.00274</td>
<td>0.0160***</td>
<td>-0.0332***</td>
<td>-0.0526***</td>
<td>-0.142***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(28.44)</td>
<td>(-0.92)</td>
<td>(-4.99)</td>
<td>(-5.40)</td>
<td>(-7.93)</td>
<td>(-19.15)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rmw2</td>
<td>-0.264***</td>
<td>-0.240***</td>
<td>-0.0171***</td>
<td></td>
<td>-0.0707***</td>
<td>0.0716***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-62.69)</td>
<td>(-50.34)</td>
<td>(-5.33)</td>
<td>(-5.40)</td>
<td>(-7.93)</td>
<td>(-19.15)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cma2</td>
<td>0.0553***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.317***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(11.35)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(27.05)</td>
<td></td>
</tr>
<tr>
<td>constan</td>
<td>0.000240***</td>
<td>0.000275***</td>
<td>0.000269***</td>
<td>-0.000253***</td>
<td>0.000807***</td>
<td>0.00114***</td>
<td>0.00111***</td>
<td>0.00120***</td>
</tr>
<tr>
<td></td>
<td>(-4.10)</td>
<td>(-4.86)</td>
<td>(-4.44)</td>
<td>(-4.62)</td>
<td>(-6.73)</td>
<td>(-9.59)</td>
<td>(-9.40)</td>
<td>(-10.16)</td>
</tr>
<tr>
<td>N</td>
<td>4901762</td>
<td>4901762</td>
<td>4901762</td>
<td>4901762</td>
<td>1002357</td>
<td>1002357</td>
<td>1002357</td>
<td>1002357</td>
</tr>
<tr>
<td>R2</td>
<td>0.186</td>
<td>0.229</td>
<td>0.230</td>
<td>0.230</td>
<td>0.0515</td>
<td>0.0713</td>
<td>0.0713</td>
<td>0.0720</td>
</tr>
</tbody>
</table>

In Table 5, the capitalization-weighted returns including dividends are presented for the market-risk premium, size factor, book-to-market factor, profitability factor, and investment pattern factor. Once again, all factors demonstrate significant explanatory power for returns at a 1% significance level. Additionally, when examining mk_rf1, smb1, and hml1, the table reveals a declining return pattern from (1) to (8). However, for cma1 and rmw1, the table indicates an increasing trend in returns from (1) to (8).

Table 6. Regression Results: Capitalization Weighted Return without Dividend

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>mk_rf2</td>
<td>1.085***</td>
<td>1.019***</td>
<td>1.014***</td>
<td>1.015***</td>
<td>0.577***</td>
<td>0.526***</td>
<td>0.522***</td>
<td>0.526***</td>
</tr>
<tr>
<td></td>
<td>(1055.89)</td>
<td>(913.82)</td>
<td>(907.22)</td>
<td>(906.58)</td>
<td>(233.38)</td>
<td>(197.80)</td>
<td>(192.38)</td>
<td>(193.70)</td>
</tr>
<tr>
<td>smb2</td>
<td>0.829***</td>
<td>0.732***</td>
<td>0.730***</td>
<td>0.509***</td>
<td>0.481***</td>
<td>0.475***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(521.83)</td>
<td>(329.73)</td>
<td>(327.08)</td>
<td>(327.08)</td>
<td>(139.82)</td>
<td>(94.66)</td>
<td>(93.44)</td>
<td></td>
</tr>
<tr>
<td>hml2</td>
<td>0.0769***</td>
<td>-0.00274</td>
<td>0.0160***</td>
<td>-0.0332***</td>
<td>-0.0526***</td>
<td>-0.142***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(28.44)</td>
<td>(-0.92)</td>
<td>(-4.99)</td>
<td>(-5.40)</td>
<td>(-7.93)</td>
<td>(-19.15)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
In Table 6, the capitalization-weighted returns excluding dividends are presented for the market-risk premium, size factor, book-to-market factor, profitability factor, and investment pattern factor. Firstly, the table reveals that all factors show significant explanatory power for returns at a 1% significance level. It is worth noting that despite the absence of dividends in the table, the results remain consistent with those including dividends. Consequently, the regression results exhibit robustness across the various terms.

### 6 CONCLUSION

We examine four widely-used asset pricing models, including the Capital Asset Pricing Model (CAMP), the Fama and French (1993) three-factor model, the Carhart (1997) four-factor model, and the Fama and French (2014) five-factor model. These inclusive five factors are Market risk factor (RM-RF), size risk factor (SMB) and Book-to-Market ratio factor (HML) Profitability Factor (RMW), and Investment Pattern Factor (CMA). We use data from CSMAR and sub-database, and the other one is China Stock Market Factor Database. By employing the data from CSMAR and sub-database, and the other one is China Stock Market Factor Database. The findings imply that the Fama and French (2014) five-factor model provides a more accurate depiction of the Chinese market compared to the Fama and French (1993) three-factor model.

### REFERENCES


The Development and Application of Enterprise Asset Lifecycle Management System Based on Big Data Technology

Jiaying Wang
1343219412@qq.com
Dongbei University of Finance and Economics, Liaoning, Dalian, 116000, China

Abstract: In the era of digital economy, the context of enterprise transformation and upgrading is becoming clearer. As an important part of realizing the whole-process digital management in enterprises, the digital construction of enterprise asset management is imperative. Therefore, this paper takes enterprise asset management as the research object, effectively integrates the application advantages of big data technology, network information technology and computer application technology, builds a data analysis and processing server relying on Hadoop cluster, and builds a digital platform for enterprise asset management based on Web by combining J2EE technical specifications. The whole platform adopts B/S architecture design, and completes the development and deployment of various functional modules and API interfaces according to MVC pattern. Aiming at many shortcomings in current enterprise asset management, the platform will introduce the whole life cycle management model, which can not only meet the requirements of enterprise asset overall planning, cost reduction and efficiency increase, but also effectively avoid the disadvantages of data island and long response period. Especially for the maintenance of assets, the system can put forward a prediction model based on C4.5 algorithm and Apriori algorithm according to the data mining theory, which is not only conducive to improving the efficiency and accuracy of asset management, but also provides a real scientific basis for enterprise asset evaluation and decision-making.

Keywords: Big Data Technology, Enterprise Asset Management, Lifecycle Management Model, Hadoop, Computer Application.

1 INTRODUCTION

The prerequisite for asset enterprises to maintain normal production and operation is also the main carrier for enterprises to realize development planning strategy. The essence of assets is the collection of various economic resources that are formed by transactions or events, owned or controlled by enterprises independently, and expected to give certain economic benefits to enterprises. Assets can be divided into current assets, fixed assets, intangible assets and other assets according to different forms. Generally, enterprise assets are mainly inventory, fixed assets and intangible assets, and directly participate in the daily production, operation and management activities of enterprises, forming financial data information, which can intuitively and comprehensively reflect the utilization effect of enterprise assets and its ability to make profits. \[1\] The value attributes of enterprise assets themselves need to be operated and deployed...
reasonably constantly, so that they can play their role to the greatest extent to reduce the risk of
depreciation, thus resulting in enterprise asset management behavior. The enterprise asset
management is an effective extension of the enterprise financial management. It is a
management tool that enterprises can realize the "real logistics" control through a perfect
management system on the basis of grasping the asset characteristics and structure scale, and it
is also a practical application efficiency.

With the development of China's economy, the management of enterprise assets has been
gradually systematized, and the scope and methods of management have been optimized and
improved, and some results have been achieved. However, there are still many problems and
shortcomings: firstly, the operating environment of enterprises is complex, and the number of
assets is increasing; Secondly, the information of asset data is frequently exchanged, and the
data distortion is serious; Thirdly, the business process lacks precise control, and the rights and
responsibilities are unequal; Fourthly, the information management means are insufficient, and
the efficiency of asset management is low. For this reason, this paper holds that enterprises
should accurately grasp the theme of high-quality development era, re-examine the strategic
development direction and the construction of internal control system, strive to promote the
digital upgrade of enterprise asset management, give full play to the application advantages of
big data technology, network information technology and computer application technology, and
build an enterprise asset management system with Hadoop cluster as data analysis and
processing server and Javaweb as interactive application system. The whole system is presented
as a Web application, which can widely support enterprise managers, financial personnel
and other business personnel to log in and use. Based on the whole life cycle management model of
the system, all business processes have been reshaped. From the perspective of overall planning,
it covers all aspects of enterprise asset planning, adding, collecting, allocating, checking,
maintaining and disposing. It overcomes the shortcomings of inconvenient user operation and
cumbersome business initiation process, effectively avoids the disadvantages of data island and
long response period, promotes the circulation and interaction of asset data, realizes resource
sharing and utilization, and provides a scientific basis for enterprise management decisions. [2]

## 2 INTRODUCTION OF KEY TECHNOLOGIES

### 2.1 Big Data Technology

The Big Data generally refers to a huge amount of data collection that can't be acquired,
managed and processed by traditional information technology and software and hardware tools
in a tolerable time. [3] The essence of big data is a huge data set, which has the characteristics of
huge information scale, diverse data types, fast update speed and low value density. The essence
of big data is a huge data set. It has the characteristics of huge information scale, diverse data
types, fast updating speed and low value density. It usually refers to a comprehensive field
dedicated to storing, analyzing and processing huge amounts of data. The big data has formed
a huge and complete ecosystem after years of development since it was put forward, which has
realized the transition from single data level to application level.

From the current technical system, big data technology is a collection of technologies covering
the whole life cycle of big data. This system covers a series of links, such as data collection,
transmission, cleaning, storage, analysis, processing, presentation and application. It provides a
complete processing paradigm for the acquisition and reflection of the value of big data, and also expands the application scenarios of big data. With the wide application of big data, the development speed of big data technology is constantly accelerating, and many technical means are gradually gathering and forming a systematic and ecological big data technology stack.

2.2 Hadoop Technical Framework

The Hadoop is an open source technical framework realized by Java language, and it is also a tool platform for distributed storage and processing of large-scale data sets. The core of Hadoop architecture is distributed file system (HDFS) and distributed computing programming framework (MapReduce), which can be used as the bottom storage and analysis core to provide users with reliable, scalable and distributed computing big data services. The Hadoop cluster deployment is supported, which is convenient for users to quickly complete the design and development of large-scale data analysis applications, and realize the parallel high-speed operation and complex call of big data. In addition, the high scalability of Hadoop framework can support the deployment and installation of various functional components, and they are compatible with each other, forming an independent application system and gradually evolving into Hadoop ecosystem. Among them, the core function of Zookeeper is to realize efficient development while maintaining distributed application coordination services; The Hive is a functional architecture based on data warehouse, which aims at mapping structured data files into a database table and converting SQL statements into MapReduce programs. As a large-scale data analysis platform, Pig can sort, filter, sum and group data through its own Pig Latin scripting language.

2.3 JavaWeb

Java is the sum total of technologies that use Java technology to solve related Web and Internet fields, and it is a comprehensive solution that uses Java language to realize dynamic Web application development. Java is subordinate to J2EE technical specification, and can use Spring+SpringMVC+Mybatis framework or Struts+Spring+Hibernate framework to complete the design and development of Web Server. The development includes functional modules, business logic control and the construction and deployment of database application model. JavaWeb usually supports B/S and C/S development modes, but B/S is more applicable, the development process is relatively concise, and the subsequent maintenance cost is low. In B/S development mode, JavaWeb can be divided into three standard layers, namely presentation layer, business logic layer and data access layer.

2.4 Data Mining

Data Mining (DM), as a kind of computer science and technology, aims at finding hidden and valuable information and knowledge from massive data. The process of data mining can be simply divided into four steps: problem definition, data preparation, data mining and result analysis. The choice of data mining method is the core of the whole data mining work, and it is also the key to build the corresponding data mining model. The common data mining methods include classification analysis, prediction analysis, cluster analysis, valuation analysis and correlation analysis, as shown in Table 1. As for the final result analysis, it is more inclined to the correlation between data, the trend and trend of data and the characteristic expression of specific objects.
Table 1: Types of Common Data Mining Algorithms.

<table>
<thead>
<tr>
<th>Classify</th>
<th>Data mining method</th>
<th>Data mining model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guided data mining</td>
<td>Classification analysis</td>
<td>Decision Tree, Random Forest, Neural Network</td>
</tr>
<tr>
<td></td>
<td>Predictive parsing analysis</td>
<td>Regression tree and rough set method</td>
</tr>
<tr>
<td></td>
<td>Valuation analysis</td>
<td>SVM, Bayesian method</td>
</tr>
<tr>
<td>Unsupervised data mining</td>
<td>Correlation analysis</td>
<td>Pearson correlation coefficient and Aprior algorithm</td>
</tr>
<tr>
<td></td>
<td>Cluster analysis</td>
<td>K-means clustering algorithm, pedigree clustering</td>
</tr>
</tbody>
</table>

2.5 Development Process

According to the application requirements of the above related application technologies, complete the configuration and deployment of the development environment of enterprise asset lifecycle management system. The development content of the system is divided into two parts. One is to build Hadoop cluster to complete the collection and storage management of all kinds of enterprise asset data. Two is to complete the development of Web Server with Spring framework under Java development environment, and complete the development and deployment of various lightweight service bus interfaces, forming a standard enterprise asset management business system.

First of all, Hadoop cluster architecture needs the support of hardware and software. The underlying operating system is Linux, CentOS 6.7(x86_64) is the version, and jdk-8u291-linux-x64 is the JDK version. According to the application requirements of the system, Hadoop will be deployed in a completely distributed cluster. There are five nodes in the cluster, named Master1, Slave1, Slave2, Slave3 and Slave4 respectively. Hadoop version is 2.7.7, which is installed in each node, and components such as Yarn, HDFS, Zookeeper, HBase and Kafka are also deployed in each node. In addition, the system will use Sqoop component to import financial data, purchasing data, warehousing data and other data into HDFS system under Hadoop framework to realize distributed storage and form original data.

Secondly, for the development of each functional module of the enterprise asset lifecycle management system, the workflow engine will be introduced to complete the process service control and performance expansion of each function. The system selects Activiti open source workflow engine, and gives the core class ProcessEngine process engine and some Service classes in Activiti to Spring container for management, so as to realize the call and processing of workflow by Web Server. [9] In addition, for the function of data mining and analysis, the system will build different data algorithm models and encapsulate them into classes that can be called directly, as shown in Figure 1, which is the key code to realize the Aprior association algorithm model.

Finally, for the development of Web application server, the basic development environment is Java, JDK version requires 1.6 or above, Java development environment is MyEclipse 2014, Web server is Tomcat 8.0, and database server is MySQL 5.5. And the project object model
Maven is used to manage the project structure. Maven chooses Apache-Maven-3.2.1 version. In the process of building the overall development environment, the installation of JDK and the configuration of environment variables are completed first to build the foundation of Java application development. Next is the installation of MyEclipse and Tomcat, the Web server, and the configuration of Tomcat is completed in the Preference option under MyEclipse. Then, based on Spring architecture, the integration and encapsulation of the whole system is completed. Through the introduction of the above key technical theories, the overall environment of system development, the configuration of related software and tools are determined, and the technical feasibility of the whole project of enterprise asset lifecycle management system is also clarified.

![Figure 1: Code for building a prior algorithm model (original).](image)

### 3 FUNCTION REALIZATION

#### 3.1 Asset Ledger

In this function module, the system will comprehensively cover the management content of various types of assets of enterprises. According to the actual situation of the system enterprise, management items such as current assets, fixed assets and intangible assets will be set up, which can truly manage the whole assets of the enterprise. The core function of asset ledger lies in the integration of asset information, that is, relying on Hadoop framework to comprehensively and timely collect and summarize enterprise asset-related data. Compared with the traditional manpower adding and uploading, the system not only improves the work efficiency, but also expands the dimension of asset information. For example, the system can display the basic information, attachment information, floor plan, financial information, purchasing information and logistics information of enterprise assets together, giving users a more intuitive feeling.

#### 3.2 Asset Operation And Maintenance

In this function module, users can initiate asset collection, transfer, inventory, maintenance, repair and other business processes. The system will preset a number of business processes, that is, define the deployment ID and name of each process through activiti.engine.ProcessEngine, and form an example that can be operated graphically. The user only needs to click to view the specific business process and the relevant person in charge.
In addition, the maintenance and repair of some fixed assets will directly affect the service life of the assets, and will also have a great impact on the daily production and operation of enterprises, so it is difficult to make a correct decision quickly. Therefore, the system provides a fault prediction model based on C4.5 decision tree and a fault correlation prediction model based on Apriori. Among them, the fault prediction model of C4.5 decision tree will complete the prediction according to the historical maintenance record data, while Apriori's fault correlation prediction model will combine the relationship between different faults and fault frequency to make inferential prediction. As shown in Table 2, the fault codes are "1" for fault occurrence and "0" for failure not occurrence. Transform the fault data into a 0-1 judgment matrix as shown in formula 1, scan and count the matrix, sum the rows and columns, and reorder it as shown in formula 2, and set the minimum support degree to 40% to get the minimum support count of 0.8. Finally, according to Apriori algorithm, the frequent itemsets of historical fault records are \{(fault 2), (fault 3, fault 4), (fault 4, fault 5)\}, and the correlation between fault codes in equipment faults can be analyzed by frequent itemsets of fault data. It can provide decision basis for subsequent operation and maintenance of equipment and asset disposal.

Table 2: Failure data record table.

<table>
<thead>
<tr>
<th></th>
<th>Fault 1</th>
<th>Fault 2</th>
<th>Fault 3</th>
<th>Fault 4</th>
<th>Fault 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maintenance record 1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Maintenance record 2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Maintenance record 3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Maintenance record 4</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

\[
M = \begin{bmatrix}
0 & 1 & 0 & 0 & 1 \\
1 & 0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 & 1
\end{bmatrix} \quad (1)
\]

\[
M' = \begin{bmatrix}
1 & 1 & 0 & 0 & 0 & 2 \\
0 & 0 & 1 & 1 & 1 & 3 \\
0 & 0 & 1 & 0 & 0 & 1 \\
1 & 1 & 0 & 1 & 0 & 3 \\
2 & 2 & 2 & 2 & 2 & 1
\end{bmatrix} \quad (2)
\]

### 3.3 Asset Disposal

In this function module, system users can select the corresponding assets in the asset list to apply for scrapping, and add the corresponding documents. The system will automatically display the specific situation of the current assets, including information such as storage location, subordinate department and specific person in charge, which is convenient for users to quickly select and complete corresponding operations.
3.4 Asset Statistics

In this function module, users can make data statistics of enterprise assets, and visually display them with graphical contents. It is beneficial for users to grasp the assets of enterprises in real time, so as to guide enterprises to make correct decision analysis. The asset statistics module includes: asset category statistics, asset change statistics, inventory statistics and service life assets statistics, etc. As shown in Figure 2, it is a statistical chart of the value of various assets.

![Figure 2: Statistical Chart of Asset Value (original).](image)

4 CONCLUSIONS

In order to promote the informatization construction of enterprise asset management, this paper integrates the application advantages of big data technology, network information technology and computer application technology efficiently, and constructs a Web-based enterprise asset management system. The system innovatively introduces the whole life cycle management model into enterprise asset management, fully considers the actual application requirements of enterprises, and comprehensively covers all aspects of the planning, addition, collection, allocation, inventory, maintenance and disposal of various enterprises’ assets. It is convenient for users to improve the processing efficiency of various businesses with simple, convenient and efficient operation, increase the value embodiment and rational application of asset data, strengthen the enterprise asset management ability, and promote the digitalization and intelligence of the internal management and control of enterprises.

REFERENCES

Evaluation of Employee Cultural Differences in Sino-US Joint Ventures
-- Based on IBM SPSS Data Analysis

Jintao Wu1,2,a, Xinlong Ding3,b*

Email: jintao.wu@lpunetwork.edu.pha
* Corresponding author: Xinlong Ding, Email: 10151@massz.edu.cnb

Department of economic and social management, Maanshan Teacher's College, Maanshan, Anhui, China1
Graduate Shool, Lyceum of the Philippines University, Cavite, Philippines2
School of Electrical and Information Engineering, Anhui University of Technology, Maanshan, Anhui, China3

Abstract: In multinational companies, employees with different cultural backgrounds because of values, beliefs, customs and language expression and so on various aspects of different, and the major problems in enterprise management in the conflict, if managers cannot rationally deal with the cultural conflict, the result is bound to cause both the opposition and conflict, if managers can make conflict fusion, the enterprise will be stable development, It is crucial to find out what cultural differences exist between employees of both sides. Firstly, this paper designed the "Cultural Difference Test Scale" in the Oriental context, and it passed the reliability and validity tests. Secondly, the cultural difference test scale was used to test the cultural differences between China and the United States, and significant differences were found. Finding out cross-cultural differences can help managers solve cross-cultural conflicts in a targeted way.

Keywords: Sino-US Joint Ventures, Cultural Differences, Cross-Cultural Management, IBM SPSS.

1 INTRODUCTION

Lv Miao & Quan Xijian (2017) believe that under the background of globalization, cultural differences in multinational enterprises will inevitably bring about cross-cultural management problems. Over the past four decades of reform and opening up, China has made remarkable achievements in its development. Its economic aggregate is now the second largest in the world. In the face of the complicated world economic situation and increasingly fierce competition in the international market, how to get out of the predicament of enterprise operation and development has become a major issue for the economic development of China and the world at large. Under the background of "Belt and Road Initiative", whether the operation and development of Chinese enterprises can accept the impact of regional differences and cultural diversity, the influence of cultural differences on management practice, and the impetus of international cultural exchanges on enterprise growth have become important research topics of management science in the future. Hofstede (2015) analyzed and compared the cultural
differences and the degree of differences among countries and regions in the world from six dimensions. Power distance (PDI), masculinity (MAS), uncertainty avoidance (UAI), individualism vs collectivism (IDV), long-term vs short-term orientation (LTO), and Indulgence vs Restraint (IVR). Patel (2013) and Geare (2014) By investigating and analyzing the values of managers, ordinary employees and college students, they have found that there are great differences among similar groups in different countries and different cultural backgrounds.

2 PROBLEMS

Farooq (2019) believes that religious belief is an important factor affecting cross-cultural management. Kivenzor (2019) believes that cultural differences in multinational enterprises will affect corporate structure and teamwork. Cultural heterogeneity is also known as cultural differences. As a result, team members often have different cognitions on management issues. Due to the cognitive differences, different managers form inconsistent opinions. Cultural difference is one of the main reasons leading to conflicts and even failures in transnational operations. Because the existing academic research results pay more attention to the decisive role of interests, psychology, environment and other factors on management decision-making, but ignore the impact of cultural differences on management decision-making. In addition, culture is recessive and difficult to quantify. This paper aims to solve the problem of measuring cross-cultural differences.

3 METHODS

IBM SPSS Statistic used in this paper is a wide range of statistical analysis software, covering scientific research, natural research, business research, policy analysis, financial analysis and other aspects. It provides a comprehensive statistical and service solution, which can not only collect, process, statistics and analysis data. Also can export beautiful chart, the data visual display. A large amount of data was generated during the investigation. IBM SPSS Statistics 19.0 software was used to process the relevant data, and independent sample test (Mann-Whitney U Test) was used to analyze the cultural differences between the two countries.

The questionnaire in this paper is based on the VSM2013 version of Hofstede's standard questionnaire on dimensions of cultural value, and combined with the Chinese context. A total of 42 questions are designed, covering six dimensions of power distance (X8, X15, X20, X23, X24, X33, X37), masculinity (X5, X13, X16, X19, X36, X39, X41), uncertainty avoidance (X1, X6, X14, X17, X18, X26, X32), individualism vs collectivism (X11, X12, X22, X25, X35, X38, X40), long-term vs short-term orientation (X2, X3, X7, X21, X28, X29, X42), and Indulgence vs Restraint (X4, X9, X10, X27, X30, X31, X34). The questionnaire is randomly distributed to 120 employees in Sino-US joint venture D. Among them, 107 are Chinese and 13 are American.

4 VALIDATION OF INSTRUMENT

The validity test mainly detects whether the questionnaire can measure the problem to be studied and reflect the extent of the problem. Generally, KMO value and Bartlett's spherical test were
used to judge the rationality of the data. Factor analysis in SPSS software was used to test the construct validity of the cultural difference test scale. The test results are shown in Table 1. It can be seen that \( \text{Chi-Sq} = 4192.016 \), significance index \( \text{Sig} = 0.000 < 0.05 \), indicating that there are common factors in the scale. It is generally believed that the KMO value should be greater than 0.5 to be suitable for factor analysis. The KMO value in the table = 0.803 > 0.5 indicates that the scale data are suitable for factor analysis and pass the validity test.

Table 1: KMO and Bartlett's tests for cultural difference factors

<table>
<thead>
<tr>
<th>KMO and Bartlett's Test</th>
<th>Kaiser-Meyer-Olkin Measure of Sampling Adequacy.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approx. Chi-Square</td>
<td>4192.016</td>
</tr>
<tr>
<td>df</td>
<td>861</td>
</tr>
<tr>
<td>Sig.</td>
<td>0.000</td>
</tr>
</tbody>
</table>

The Cronbach a coefficient is usually used to test the reliability of the questionnaire. If the coefficient is greater than or equal to 0.8, it indicates that the questionnaire can reflect the problem well and the reliability is high, but if the coefficient is below 0.6, it indicates that the questionnaire has problems. In order to obtain better results, the problem should be adjusted. We using SPSS software for internal consistency of the questionnaire were tested, the questionnaire using the level 5 scale suitable for cronbach a coefficient test, test results such as table 2, it can be seen that the employees completed the questionnaire cronbach a coefficient test were greater than 0.8, with internal consistency, suggests that cultural differences test scale has high reliability.

Table 2: Reliability Statistic

<table>
<thead>
<tr>
<th>The data source</th>
<th>Cronbach’s Alpha</th>
<th>N of Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chinese data</td>
<td>0.947</td>
<td>42</td>
</tr>
<tr>
<td>U.S. data</td>
<td>0.909</td>
<td>42</td>
</tr>
<tr>
<td>Combined data</td>
<td>0.941</td>
<td>42</td>
</tr>
</tbody>
</table>

5 RESULTS

5.1 Hypothesis 1 (H1): There are cultural differences between Chinese employees and American employees in enterprise D.

Based on the data collected by questionnaire survey of employees in D enterprise, Mann-Whitney U test was used to test whether there were differences between Chinese and American employees in 42 cultural variables (X1~X42), and hypotheses H1(1) ~ H1(42) were established. with sig. < 0.05, which was significant, and the null hypothesis was rejected. Sig > 0.05, the test is not significant, and the null hypothesis is not rejected, indicating that there is no sufficient evidence to show that there is a difference between Chinese and American cultural variables or that there is no difference in a cultural variable, and the mean difference between Chinese and American employees in various cultural variables is calculated, the mean difference is > 0.5 or
< -0.5, which is considered to be a big difference. This is consistent with the results of Mann-Whitney U test. The data were arranged as shown in Table 3.

Table 3 Hypothesis Test Summary (1)

<table>
<thead>
<tr>
<th>Item</th>
<th>Null Hypothesis</th>
<th>sig.</th>
<th>Decision</th>
<th>Mean</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The distribution of Personal health is the same across categories of Q2_Nationality.</td>
<td>0.982</td>
<td>Retain the null hypothesis.</td>
<td>-0.040</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>The distribution of frugal is the same across categories of Q2_Nationality.</td>
<td>0.000</td>
<td>Reject the null hypothesis.</td>
<td>1.493</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>The distribution of Unwavering effort is the same across categories of Q2_Nationality.</td>
<td>0.603</td>
<td>Retain the null hypothesis.</td>
<td>-0.168</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>The distribution of Emotional expression is the same across categories of Q2_Nationality.</td>
<td>0.939</td>
<td>Retain the null hypothesis.</td>
<td>-0.080</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>The distribution of Harmony between man and nature is the same across categories of Q2_Nationality.</td>
<td>0.084</td>
<td>Retain the null hypothesis.</td>
<td>-0.0565</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>The distribution of Accuracy and punctuality is the same across categories of Q2_Nationality.</td>
<td>0.263</td>
<td>Retain the null hypothesis.</td>
<td>-0.336</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>The distribution of Helping friends is the same across categories of Q2_Nationality.</td>
<td>0.000</td>
<td>Reject the null hypothesis.</td>
<td>1.279</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>The distribution of The decision of the supervisor involves your work and should consult your opinion is the same across categories of Q2_Nationality.</td>
<td>0.574</td>
<td>Retain the null hypothesis.</td>
<td>0.111</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>The distribution of Private living space is the same across categories of Q2_Nationality.</td>
<td>0.533</td>
<td>Retain the null hypothesis.</td>
<td>-0.211</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>The distribution of Joyfulness is the same across categories of Q2_Nationality.</td>
<td>0.940</td>
<td>Retain the null hypothesis.</td>
<td>0.023</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>The distribution of Focus on your Image is the same across categories of Q2_Nationality.</td>
<td>0.227</td>
<td>Retain the null hypothesis.</td>
<td>0.372</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>The distribution of Your work is respected by your family and friends is the same across categories of Q2_Nationality.</td>
<td>0.741</td>
<td>Retain the null hypothesis.</td>
<td>-0.106</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>The distribution of Most people can be trusted is the same across categories of Q2_Nationality.</td>
<td>0.411</td>
<td>Retain the null hypothesis.</td>
<td>0.188</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>The distribution of A good manager does not give precise answers to most of the questions his subordinates encounter and ask at work is the same across categories of Q2_Nationality.</td>
<td>0.000</td>
<td>Reject the null hypothesis.</td>
<td>1.246</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Statement</td>
<td>p-value</td>
<td>Action</td>
<td>z-value</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---------------------------------------------------------------------------</td>
<td>----------</td>
<td>-----------------------------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>The distribution of Organizational structures should avoid having two bosses for one subordinate is the same across categories of Q2_Nationality.</td>
<td>0.619</td>
<td>Retain the null hypothesis.</td>
<td>-0.197</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>The distribution of Competition is the same across categories of Q2_Nationality.</td>
<td>0.811</td>
<td>Retain the null hypothesis.</td>
<td>0.047</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>The distribution of There should be flexibility in organizational regulations is the same across categories of Q2_Nationality.</td>
<td>0.000</td>
<td>Reject the null hypothesis.</td>
<td>1.442</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>The distribution of Uncertain events should be avoided is the same across categories of Q2_Nationality.</td>
<td>0.007</td>
<td>Reject the null hypothesis.</td>
<td>-0.792</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>The distribution of Give recognition for good performance is the same across categories of Q2_Nationality.</td>
<td>0.882</td>
<td>Retain the null hypothesis.</td>
<td>0.018</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>The distribution of Power at the top comes from position is the same across categories of Q2_Nationality.</td>
<td>0.982</td>
<td>Retain the null hypothesis.</td>
<td>-0.024</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>The distribution of National pride is the same across categories of Q2_Nationality.</td>
<td>0.691</td>
<td>Retain the null hypothesis.</td>
<td>-0.192</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>The distribution of Everyone should pursue his own best interests is the same across categories of Q2_Nationality.</td>
<td>0.000</td>
<td>Reject the null hypothesis.</td>
<td>-1.370</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>The distribution of Relatives do not have priority in employment is the same across categories of Q2_Nationality.</td>
<td>0.848</td>
<td>Retain the null hypothesis.</td>
<td>-0.117</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>The distribution of Confrontation is normal is the same across categories of Q2_Nationality.</td>
<td>0.144</td>
<td>Retain the null hypothesis.</td>
<td>-0.379</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>The distribution of The interest of the individual precedes that of the group is the same across categories of Q2_Nationality.</td>
<td>0.000</td>
<td>Reject the null hypothesis.</td>
<td>-1.262</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>The distribution of Multiple religions can coexist is the same across categories of Q2_Nationality.</td>
<td>0.283</td>
<td>Retain the null hypothesis.</td>
<td>0.302</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>The distribution of What you really want to do often gets in the way of others and outsiders is the same across categories of Q2_Nationality.</td>
<td>0.570</td>
<td>Retain the null hypothesis.</td>
<td>-0.143</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>The distribution of The most important events in life are in the future, not in the past or the present is the same across categories of Q2_Nationality.</td>
<td>0.002</td>
<td>Reject the null hypothesis.</td>
<td>1.004</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>The distribution of Save most of your income is the same across categories of Q2_Nationality.</td>
<td>0.000</td>
<td>Reject the null hypothesis.</td>
<td>1.197</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>The distribution of Even without love, there are family responsibilities is the same across categories of Q2_Nationality.</td>
<td>0.488</td>
<td>Retain the null hypothesis.</td>
<td>-0.232</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>The distribution of Instant gratification when a need is met is the same across categories of Q2_Nationality.</td>
<td>0.942</td>
<td>Retain the null hypothesis.</td>
<td>0.032</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>The distribution of You often feel nervous, stressed or anxious is the same across categories of Q2_Nationality.</td>
<td>0.000</td>
<td>Reject the null hypothesis.</td>
<td>1.285</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>The distribution of Fear of arguing with superiors or teachers is the same across categories of Q2_Nationality.</td>
<td>0.177</td>
<td>Retain the null hypothesis.</td>
<td>-0.461</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>The distribution of You feel humiliated for making a mistake is the same across categories of Q2_Nationality.</td>
<td>0.001</td>
<td>Reject the null hypothesis.</td>
<td>0.931</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>The distribution of Leave plenty of time for yourself and your family after work is the same across categories of Q2_Nationality.</td>
<td>0.006</td>
<td>Reject the null hypothesis.</td>
<td>-1.009</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>The distribution of Good working environment is the same across categories of Q2_Nationality.</td>
<td>0.937</td>
<td>Retain the null hypothesis.</td>
<td>-0.041</td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>The distribution of Maintain a good working relationship with your immediate supervisor is the same across categories of Q2_Nationality.</td>
<td>0.773</td>
<td>Retain the null hypothesis.</td>
<td>0.134</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>The distribution of Employees have job security is the same across categories of Q2_Nationality.</td>
<td>0.950</td>
<td>Retain the null hypothesis.</td>
<td>0.011</td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>The distribution of Work with friendly people is the same across categories of Q2_Nationality.</td>
<td>0.002</td>
<td>Reject the null hypothesis.</td>
<td>-0.362</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>The distribution of Interesting work content is the same across categories of Q2_Nationality.</td>
<td>0.450</td>
<td>Retain the null hypothesis.</td>
<td>-0.190</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>The distribution of Have an opportunity for advancement is the same across categories of Q2_Nationality.</td>
<td>0.000</td>
<td>Reject the null hypothesis.</td>
<td>-1.168</td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>The distribution of The work is diverse and adventurous is the same across categories of Q2_Nationality.</td>
<td>0.000</td>
<td>Reject the null hypothesis.</td>
<td>-1.168</td>
<td></td>
</tr>
</tbody>
</table>

Asymptotic significances are displayed. The significance level is 0.05. Note: the data is compiled by the SPSS operation results, and see the output results of the data.

By the above data can see, in D enterprise, Sino-US cultural differences significant items as follows: X2, X7, X11, X14, X17, X18, X22, X25, X26, X28, X29, X32, X34, X35, X40 and X42, a total of 16. According to Hofstede's cultural dimension framework, among the 16 cultural variables with significant differences, X11, X22, X25, X35 and X40 belong to the cultural dimension of individualism/collectivism. Chinese employees show collectivist value orientation, while American employees show individualistic value orientation. X11, X22, X25,
X35 and X40 belong to the cultural dimension of uncertainty avoidance, Chinese employees show a low level of uncertainty avoidance, American employees show a high level of uncertainty avoidance, X2, X7, X28, X29 and X42 belong to the cultural dimension of long-term and short-term orientation, Chinese employees show a long-term orientation, American employees show a short-term orientation. In addition, among the variables of indulgence and restraint dimension, only X34 "you feel humiliated because you made a mistake" has significant difference, while the other 25 cultural variables have no significant difference, namely, the results of Distance power dimension, masculinity dimension and indulgence and restraint dimension are different from those of Hofstede’s test.

5.2 Hypothesis 2 (H2) : Cultural differences lead to cross-cultural management conflicts.

According to the data summary, Mann-Whitney U test was used to test whether there were differences or conflicts between Chinese and American employees in 9 cross-cultural management behaviors (Y1~Y9). The hypothesis H2 (1) ~H2 (9) was established, and the Sig. < 0.05 was significant, and the null hypothesis was rejected, indicating that the cultural difference between China and the United States would bring about cross-cultural management conflicts. The Sig. > 0.05 was not significant, and the null hypothesis was not rejected, indicating that there was no sufficient evidence that the cultural difference between China and the United States would bring about cross-cultural management conflicts, or the opinions were the same. The data are summarized as shown in Table 4.

<table>
<thead>
<tr>
<th>Item</th>
<th>Null Hypothesis</th>
<th>Sig.</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The distribution of Whether you accept the new management concept and mode is the same across categories of Q2_Nationality.</td>
<td>0.002</td>
<td>Reject the null hypothesis.</td>
</tr>
<tr>
<td>2</td>
<td>The distribution of Do you prefer rational or empirical decision making is the same across categories of Q2_Nationality.</td>
<td>0.992</td>
<td>Retain the null hypothesis.</td>
</tr>
<tr>
<td>3</td>
<td>The distribution of Do you think interpersonal relationships have an impact on the success of a task is the same across categories of Q2_Nationality.</td>
<td>0.006</td>
<td>Reject the null hypothesis.</td>
</tr>
<tr>
<td>4</td>
<td>The distribution of Whether you will accept competitive enterprises to employ high salaries is the same across categories of Q2_Nationality.</td>
<td>0.770</td>
<td>Retain the null hypothesis.</td>
</tr>
<tr>
<td>5</td>
<td>The distribution of Do you prefer to work for a company with a sense of belonging or one that focuses on personal achievement is the same across categories of Q2_Nationality.</td>
<td>0.003</td>
<td>Reject the null hypothesis.</td>
</tr>
<tr>
<td>6</td>
<td>The distribution of Are you willing to spend a lot of money to get your employees to work more efficiently is the same across categories of Q2_Nationality.</td>
<td>0.770</td>
<td>Retain the null hypothesis.</td>
</tr>
</tbody>
</table>
same across categories of Q2_Nationality.

<table>
<thead>
<tr>
<th></th>
<th>The distribution of</th>
<th></th>
<th>The distribution of</th>
<th></th>
<th>The distribution of</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>Are you willing to participate in the activities organized by employees is the same across categories of Q2_Nationality.</td>
<td>0.893</td>
<td>Retain the null hypothesis.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>If the company assigns you to a branch abroad, will you make adequate preparation is the same across categories of Q2_Nationality.</td>
<td>0.706</td>
<td>Retain the null hypothesis.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>When making decisions, do you take the plunge or do you hold a meeting first is the same across categories of Q2_Nationality.</td>
<td>0.006</td>
<td>Reject the null hypothesis.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Asymptotic significances are displayed. The significance level is 0.05. Note: the data is compiled by the SPSS operation results, and see the output results of the data.

It can be seen from the above data that Y1, Y3, Y5 and Y9 are significant in the variables of cross-cultural management conflict, indicating that Chinese and American employees in enterprise D have conflicts in Y1, Y3, Y5 and Y9. It can be concluded that the cultural differences between China and America will lead to conflicts in management concepts, interpersonal relationship management, human resource management and leadership style and decision-making. The remaining five items show the same views. The influence of significantly different cultural variables on cross-cultural management conflict will be analyzed in detail in the next section.

6 CONCLUSIONS

The Mann-Whitney U Test method was used to conduct statistical tests on the cultural difference test scale filled in by Chinese and American employees; The three dimensions of UAI, LTO and IDV were found to have significant differences between Chinese and American cultural variables; Through descriptive statistical analysis, the specific manifestation of Sino-US cultural differences is expounded in detail.

In terms of the cultural dimension of UAI, the differences between Chinese and American cultures are mainly as follows: Chinese employees can adapt to unstructured situations, do not stick to the format or rules, can change the way of doing things according to people and events, can be positive and optimistic to deal with the uncertain things, and do not complain about the superiors do not have accurate instructions; Chinese employees are more tolerant and expect success, so they will feel pressure and tension. American employees have a lower tolerance to the uncertain environment and feel uncomfortable and unadaptable in the unstructured environment, meanwhile, they make more efforts to prevent the occurrence of unstructured situations, such as paying more attention to rules and regulations; Focus on precision in solving problems, usually requiring explicit instructions from superiors; Expect success but not afraid of failure, so personal pressure and tension are low.
In the cultural dimension of IDV, the differences between Chinese and American cultures are mainly reflected in the following aspects: Chinese employees believe that collective interests are higher than individual interests; Collectivism tendency, pay more attention to the feelings of others in the organization, tend to cooperate; At the same time, personal and family time will be sacrificed for work because of a strong sense of community honor. American employees tend to prefer individual interests to group interests; They value personal and family above work and are unwilling to sacrifice leisure time for work, more willing to take risks alone, like to have fun work, like to get a sense of achievement from competition.

In the cultural dimension of LTO, the cultural differences between China and the United States are mainly reflected in the following aspects: Chinese employees pay attention to thrift and perseverance, personal adaptability, relative values, they good at coordination, consider present and future as a whole and pay more attention to the future. American employees pay attention to personal interests, focus on the present and expect quick results; Embrace absolute values and separate home and work completely, focus on social consumption.

**Funding:** Youth Project of Anhui Province, Grant/Award Number: gxyqZD2022105

**REFERENCES**


Discuss the Financial Strategy Management of Technology Companies

Meiying Cheng

e-mail: 1216930130@qq.com

Guangdong Neusoft University, Nanhai Software Science Park, Foshan, Guangdong, China

Abstract: An important part of the enterprise development strategy is the financial strategy. If you want to make the enterprise develop steadily, you need to carry out scientific strategies, especially financial strategic planning. The operation of an enterprise is a dynamic change process, with its own life cycle, and the characteristics of each cycle are different. In response to these differences, the specific methods of each stage will also change from time to time. In the face of many complex changes, not only do large textile enterprises need to cope with the pressure brought by risks, but a large number of small and medium-sized textile enterprises in the industry also need to improve their strength and reduce the impact of risks. Against this background, Technology companies, as a technology enterprise, determines the current cycle of the company, adjusts its financial strategy according to the characteristics of the life cycle, and reasonably plans the financial content of the enterprise, which is of great significance to its long-term development.


1 INTRODUCTION

Financial strategy is an objective requirement for financial management activities under the new era. It integrates the principles and methods of strategic management. Based on adapting to the environment and scientific use of existing conditions, it aims to achieve the long-term balanced flow and allocation of enterprise funds, so that enterprises have strong financial capabilities, so as to carry out long-term planning and management of enterprise financial activities. For enterprises, if they want to develop continuously and healthy, they must have the ability to continuously create value and maximize value. To achieve this, we need to take value creation as the orientation and implement value management in the financial strategic management link. This article takes this perspective as the starting point, comprehensively applies many theories, such as budget, value, strategy, finance and other management theories, and performance evaluation theory, takes value creation as the core, further expands the scope of financial strategic management, significantly improves the value creation ability of enterprises, and then promotes enterprises to achieve the goal of maximizing value.. This paper studies the financial strategy management of enterprises based on value creation[1].

1.1 Research Background

Technology companies, It belongs to a local high-tech enterprise. The company is mainly responsible for the industry of science and technology, chemical promotion, and the operation
of science and technology products and services. At the same time, it also takes into account the research and development of some technical products. In the process of the industrialization of the company's high technology, all management activities of the enterprise must be based on strategic management. The main contents of enterprise strategic management include the management of future uncertainties. According to the characteristics and needs of technology enterprises, it is necessary to study the management theory and management methods of advanced science and technology enterprises, provide theoretical support for the development of enterprises, and enhance the core competitiveness of enterprises. Financial strategy is very important for an enterprise. It stands in the future to advise enterprises and formulate strategies suitable for the enterprise itself and make it sustainable. Enterprises should strengthen financial strategic management, do a good job in long-term planning, refine the planning into each year, and plan the annual financial budget goals. Strengthening the financial strategic management of enterprises is very helpful for enterprises to achieve their business goals. The catalytic role.

1.2 Research Significance

The development of e-commerce is as fast as sitting on a rocket. The e-commerce industry is developing rapidly, and the application of e-commerce in all aspects is constantly broadening and deepening.

At present, with the continuous development and growth of e-commerce enterprises such as Taobao Tmall, JD.com and Alibaba, the volume of electronic transactions has also increased rapidly. At the same time, it has also affected the current traditional financial management environment, so the development of emerging e-commerce financial management is imminent. Strategic financial management is summarized from people's experience in practice. The specific theoretical basis is based on a certain logical relationship. Organic combination is used for guidance. Financial management is a complete and multi-level theoretical system of practice, strategic financial management includes the following:

1) Strategic planning. Strategic financial management has been demarcated. Enterprise body Tax Mission, formulate Guiding principle, determine Target And set the goals of the financial management subject. Plan.

2) Strategic decision making. Evaluate the strategic management program and choose the most suitable one. Corporate strategy Develop survival programs.

3) Strategic control. Determine the implementation strategy Organizational structure, Leadership incentive and monitoring strategies in the main goal of financial management. Effectiveness.

4) Strategic performance evaluation. Performance evaluation is between the activities of enterprise members and the enterprise strategy. Bond It is motivation, evaluation and control. Staff Behaviour An effective tool. Performance evaluation usually includes Performance indicators Setting, scale, evaluation, control, feedback, adjustment, motivation, etc. Strategic
performance evaluation revolves around the Strategic management Goal to carry out and promote Business objectives of the enterprise. It was successfully completed.

2 FINANCIAL STRATEGIC MANAGEMENT ANALYSIS OF SCIENCE AND TECHNOLOGY ENTERPRISES

2.1 Issue Statement

Traditional financial management is no longer the development of e-commerce. The development of e-commerce has led to the emergence of new e-commerce accounting. This question will analyze the problems faced by e-commerce enterprises and study the difference between traditional accounting and emerging e-commerce accounting. In the end, the financial strategic management chosen by each enterprise allows e-commerce enterprises to develop continuously and steadily. Accounting for e-commerce will change the original model and form, improve the quality of accounting work, and also promote the development of accounting. Researchers chose this topic to learn the information technology financial war. The effectiveness of minor management. Through analysis, the following problems will be solved:

1. Analyzing the financial management environment and combining the requirements of the overall strategy of the enterprise, it improves the financial ability of the enterprise.
2. Financial strategy management focuses on the systematic analysis of financial statements, which improves the overall coordination of the enterprise;
3. Financial strategic management focuses on long-term interests and overall performance, which helps to create and maintain the financial management advantages of the enterprise, thus creating and maintaining the competitive advantage of the enterprise; 
5. All advantages and disadvantages can be clearly seen through data analysis;
6. Collect financial data and establish more reasonable strategic goals;
7. Observe the financial strategic development plan and profit-centered financial plan;
8. Determine the effectiveness of financial strategy management.
9. It can solve the practical problems existing in the enterprise.
10. Improve the operating efficiency of the enterprise and maximize the value of the enterprise.
11. Try to use cost-effective methods to quickly solve problems arising in implementation.

2.2 Research and Design

Content determines the form, and the nature of social sciences determines the choice of research methods. Like natural sciences, using strict mathematical statistical methods to compile human factors into numbers, and information will be missing (Li Gaoyong and Mao Jiye, 2015). Compared with other methods, management qualitative research with case studies as the main content is easy to explore the significance and purpose of human participation. Due to the theoretical framework of the relationship between business model, financial strategy and enterprise value, and based on the business model and financial strategy, effective measurement
indicators have not yet been formed, so that the empirical analysis method based on large samples cannot be used. Therefore, this paper uses a qualitative research method based on case studies\[3\].

In recent years Technology companies A lot of money has been invested in product technology research and development. In the next few years, the company will still strive to complete the project establishment of provincial and municipal science and Technology committees in a way, so that Technology companies Become the leader in the era of the development of China's IoT. Technology companies Long-term development is being achieved in the fierce industry market competition with a prudent overall strategy, and the market share is gradually expanding. Actively cater to the development trend of the domestic and foreign economic environment under the new normal, actively grasp the adjustment of the national industrial structure, adjust the company's industrial structure, coordinate the relationship between R&D projects and the market, and enable the traditional remote water meter products of enterprises to obtain new vitality with the help of new technologies.

2.3 Financial Strategy of Technology Companies.

2.3.1 Investment Strategy\[4\]

In recent years, the company has continuously promoted the R&D process of independent brands. While expanding the profit level, it optimizes the allocation of resources by improving R&D investment. In recent years Technology companies While increasing R&D investment, a series of policies have also been implemented to control the input-output efficiency of R&D support. On the basis of improving R&D efficiency, a certain R&D cost will be reduced in 2020. Objectively speaking, the enterprise investment strategy mainly includes two aspects: internal investment strategy and external investment strategy. Technology companies Investment in fixed assets and intangible assets is internal investment, and long-term investment is foreign investment.

While increasing R&D investment, the company also implemented a series of policies to control the input-output efficiency of R&D expenditure. On the basis of improving R&D efficiency, certain R&D costs were reduced in 2020. It can be seen from Table 1 that enterprise investment strategy mainly includes internal investment strategy and external investment strategy. A Technology Co., Ltd.'s investment in fixed assets and intangible assets belongs to internal investment, while long-term investment belongs to external investment\[5\].

| Table 1 A Investment of Science and Technology Co., Ltd. from 2016 to 2018 Unit: yuan |
|---------------------------------|-------------------|--|-------------------|
|                                 | 2018\(^{\text{year}}\) | 2019\(^{\text{year}}\) | 2020\(^{\text{year}}\) |
| fixed assets                    | 90,999,650.90      | 76,260,545.25      | 74,337,207.22      |
| Intangible assets               | 8,107,796.18       | 8,056,044.14       | 9,046,515.98       |
| long-term investment            | -                 | -                 | 11,343,000.00      |
| Total assets                    | 159,891,917.32     | 182,449,115.30     | 196,596,556.04     |
| Fixed assets/total assets (%)   | 56.91              | 41.8              | 37.81              |
| Intangible assets/total assets (%) | 5.07           | 4.42              | 4.6                |
| Long term investment/total assets (%) |          |                    | 5.77               |

Note: The data are sorted out according to the audit report of Science and Technology Co., Ltd. from 2018 to 2020.
2.3.2 Analysis of the Causes of the Financial Strategy of Technology Companies.

1. Lack of reasonable evaluation of corporate investment
2. Conservative fundraising strategies lead to unreasonable funding structures.
3. Lack of effective incentives for income distribution

2.3.3 Evaluation of the Effectiveness of the Implementation of the Financial Strategy

The impact of each branch strategy in the financial strategy on the enterprise's financial strategy varies, so there are also differences in the importance of each indicator set according to the enterprise's financial strategic plan, so it is necessary to distribute the weight of the selected relevant indicators and determine their importance in the financial strategy. When assigning weights to indicators, this paper chooses gray correlation analysis to escort the applicability and operability of subsequent calculations and evaluations.

Technology companies

The dividend distribution strategy is also unsatisfactory, and the score is only 3.4522,

As the strategy with the lowest score in financial strategy, it has many disadvantages:

Technology companies. The interests of shareholders have not been fully considered. Although the high dividend distribution policy can improve the image of enterprises among investors, the cost of capital will increase with the increase of the number of shares, and Technology companies

The current business status does not match. and

Through Table 1, we can analyze technology companies. If a higher dividend is distributed to shareholders, it will reduce the retained earnings of the enterprise and indirectly reduce the amount of internal financing of the enterprise. Technology companies. External financing is required, which will generate financing costs and increase. Technology companies, The cost is difficult to maintain. Technology companies, Continuous production and operation.

![Financial Strategy Matrix](image)

**Figure1** Financial Strategy Matrix
3 TECHNOLOGY COMPANY

Suggestions for improving the fundraising strategy

3.1 Propose

According to the company's overall strategic goals, the current maturity period of technology companies is necessary to maintain stability, gradually expand production and consolidate the market position. In order to maintain this stability, its strategic focus is to improve efficiency. At the same time, it also takes into account the requirements of improving sales level and service quality. The financial strategy of maturity needs to achieve sustainable development, optimize the allocation of funds, and maximize the benefits of capital.

3.2 Technology Companies suggestions for Improving the Fundraising Strategy

1. Broaden external financing channels

The company's financing channel is single, and more financial support may be needed if it wants to have new developments in the future. At present, the company's tendency in debt financing is weak, and the debt ratio is low. Therefore, in the long run, in order to meet the needs of the company to continue to expand its scale, financing should be more active. Consider increasing debt financing for allocation, with debt financing as the main external financing channel.

In the financing process, companies should focus on solving financial risk control problems as a key part. According to the actual situation, determine the appropriate financing rate and strive to achieve the goal at a lower cost, so that the company can raise enough funds.

Technology companies too much attention is paid to domestic investment in the direction of investment, but the intensity of foreign investment is small. Although domestic investment can improve the production efficiency of enterprises and expand the market size. You can't rush to invest in areas that are too different from the company. You can take the market as the unified core and reduce investment mistakes.

2. Improve core competitiveness

Core competitiveness affects the survival of enterprises and is the core of enterprise development. Enterprises should not only expand their production scale, but also tend to improve their core competitiveness and cultivate and expand their own advantages, so as to win the favor of more buyers in the protective clothing market. At the same time, it is necessary to strengthen R&D and marketing, and expand brand influence. This can form a brand effect, attract more customers, promote a steady increase in sales, and make more profits.

3.3 Technology Companies suggestions for Improving the Profit Distribution Strategy

1. Establish a sound incentive distribution mechanism

2. Improve financial performance evaluation

During the actual implementation of the financial strategy, there may be many situations that cause optimization measures to fail to achieve the purpose of optimization. In order to avoid this situation, in the process of optimizing the financial strategy, the following five safeguards
are formulated to ensure that technology companies' financial strategies have been effectively implemented.

3.4 Financial Strategic Safeguard Measures of Technology Companies

If the financial strategy is to be effectively implemented, the preparation before implementation is indispensable. The perfection of the preliminary preparation will directly affect the overall implementation of the financial strategy. Through hierarchical management and supervision of budgeting, the rationality of budgeting is ensured. The decomposition and implementation of the budget also decomposes the goals and responsibilities, thus effectively ensuring the implementation of the financial strategy and gradually achieving the overall goals of the enterprise. The specific programming cycle is shown in Figure.

3.5 Establish an Evaluation and Incentive Mechanism for Financial Strategy Implementation

After making the preliminary preparations, in order to ensure the continuation of the financial strategy work, it is also necessary to pay attention to the assessment and evaluation of the financial strategy within a certain period of time and the motivation of relevant personnel afterwards.

1. Establish a financial strategy implementation assessment and evaluation system
2. Combining performance and incentive mechanism

3.6 Strengthen Management's Awareness of Financial Strategy

Whether the financial strategy can be effectively implemented is closely related to management's financial strategy awareness. The implementation of financial strategies cannot rely solely on the financial department, but requires collaboration and coordination among enterprises as a whole. For example, in terms of departmental budgets, if enterprises cannot reach a unified understanding, departments are likely to focus on their own interests and do not consider the company's overall strategic implementation, resulting in the implementation of financial strategies being formal and cannot be truly implemented. Based on this situation, the establishment of managers' financial strategic awareness is a prerequisite for the effective implementation of the financial strategy of the whole enterprise.

Strengthening risk control; and establishing an information feedback system to achieve the long-term and stable development of enterprises.
3.7 Analysis of Business Risk Environment

Operational risk analysis includes the overall environment and industry environment. The combination of commercial risk and financial risk has four characteristics. Figure 2. To analyze the overall environmental factors, the key is to determine the opportunities and threats of enterprises. At present, the overall environmental factors can be divided into: first, economic factors. The characteristics and development direction of the relevant competitive economy that enterprises need to participate in is the economic environment. It not only affects the growth of enterprise value, but also affects the formation of enterprise financial strategy. Second, finance. Under the market economy environment, the financing and investment of enterprises cannot be separated from the support of the financial environment. Type and quantity of financial institutions and related financial products, and the maturity of financial markets will have an impact on the capital of enterprises, thereby affecting their value. Third, the market size. The larger the market size, the greater the possibility of segmentation. Therefore, both large and small enterprises believe that there is a greater opportunity to scale up. Fourth, laws and policies. National macro laws and policies will have a significant impact on the competitiveness of enterprises. Fifth, technology, science and technology are undoubtedly the primary productive forces. The technical level will directly affect the scientificity and efficiency of production. In addition, it can also provide guidance for the investment direction of enterprise funds. Sixth, social culture. The changing trend of public life patterns and values is very important for consumer goods manufacturers. The demand for industrial goods often comes from the consumer goods industry, that is, the dynamic change of the demand for consumer goods industry will significantly affect the demand for industrial goods. Therefore, the dynamic changes of social and cultural factors will have an impact on many industries.
3.8 Profitability Analysis

For real enterprises, profitability analysis usually takes indicators such as return on equity and return on total assets, which can be sorted out.

Table 2. Profitability Analysis Table of Guangdong A Information Technology Co., LTD  Unit: 

<table>
<thead>
<tr>
<th></th>
<th>In 2018,</th>
<th>In 2019,</th>
<th>In 2020,</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Total assets</td>
<td>113473911.2</td>
<td>171170516.3</td>
<td>189522835.7</td>
</tr>
<tr>
<td>Return on equity (%)</td>
<td>13.97</td>
<td>8.42</td>
<td>8.69</td>
</tr>
<tr>
<td>Return on Total Assets (%)</td>
<td>17.09</td>
<td>7.46</td>
<td>7.62</td>
</tr>
</tbody>
</table>

Note: The data are compiled according to the 2018-2020 audit report of Guangdong A Information Technology Co., LTD.

It can be seen from Table 2 that the profitability analysis of Guangdong A Information Technology Co., Ltd. from 2018 to 2020 shows that the average total assets of Guangdong A Information Technology Co., Ltd. has continued to rise, but the return on net assets and the return on total assets have both declined, indicating that the company's profits have declined, but its main business capacity is still relatively competitive compared with other water meter industries in China.

3.9 Operational Capacity Analysis

As can be seen from the operating capacity analysis, the accounts receivable turnover ratio of A Information Technology Co., Ltd. was maintained between 3.5 and 2 from 2018 to 2019, which indicates that the company's accounts receivable recovery is not ideal, and the company's fund use efficiency is relatively cross. However, the current asset turnover ratio of Guangdong A Information Technology Co., Ltd. between 2018 and 2019 was greater than 1, indicating that the company's main business income was higher than the current assets at this stage, the enterprise's current asset turnover rate was also fast, and the utilization of current assets was relatively ideal. However, the turnover of current assets dropped to 0.84 in 2018, indicating that the turnover of current assets of the company declined in 2020, and the utilization of current assets needs to be improved. From 2018 to 2020, the total asset turnover of the company did not exceed 0.6, which was not high compared with the same industry, indicating that the investment efficiency of the company was not good.

Table 3 Operational Capacity Analysis Table of Guangdong A Information Technology Co., LTD  Unit: time

<table>
<thead>
<tr>
<th></th>
<th>In 2018,</th>
<th>In 2019,</th>
<th>In 2020,</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accounts receivable turnover</td>
<td>3.29</td>
<td>3.24</td>
<td>2.39</td>
</tr>
<tr>
<td>Current asset turnover</td>
<td>1.46</td>
<td>1.15</td>
<td>0.84</td>
</tr>
<tr>
<td>Total asset turnover</td>
<td>0.59</td>
<td>0.53</td>
<td>0.44</td>
</tr>
</tbody>
</table>

Note: The data are compiled according to the 2018-2020 audit report of Guangdong A Information Technology Co., LTD.
3.10 Comprehensive Analysis of Financial Position

By analyzing the financial statements of Guangdong A Information Technology Co., LTD from 2018 to 2020, the current financial strategy of the company makes the company's financial status in terms of debt paying ability, profitability and operating ability as follows:

First, A information technology co., LTD. Guangdong asset-liability ratio reached the highest in 2019, mainly due to the influence of the new financing, this phase of the financial leverage ratio is relatively low, the specific performance in the short-term debt paying ability strong, but its debt management and financing model has relatively large limitation, and poor liquidity, However, the asset-liability ratio declined again in 2020.

Second, the return on assets of Guangdong A Information Technology Co., Ltd. showed a slow downward trend. After 2019, the company began to try external debt financing and the financial expenses also affected the overall business performance of the company.

Third, Guangdong A Information Technology Co., LTD., from 2018 to 2020, the company's capital use efficiency was high, and the current asset turnover rate in 2018 and 2019 remained above 1, indicating that the current asset utilization rate at this stage was in line with the industry average level, but the current asset turnover rate decreased in 2020.

In conclusion, before 2018, Guangdong A Information Technology Co., Ltd. mainly used internal financing. In 2019, the enterprise began to try external financing, but the financing mode was single and the scale was small. The decline of fixed asset investment has not effectively improved the efficiency of enterprise asset investment. At present, the former Guangdong A Information Technology Co., Ltd. relies solely on the financial budget management strategy, which leads to the insufficient growth of the enterprise's financial control ability, and it is urgent to optimize the financial strategy.

3.11 Establish A High-Level Risk Management Awareness of Enterprises

The outbreak of COVID-19 is also Technology companies. It brings opportunities and challenges. Affected by the COVID-19 pandemic, physical stores have been basically closed, which has a huge impact on enterprises dominated by offline retail channels, but Technology companies. The main sales channel is online channels, right. Technology companies. The impact is less. At the same time, with the closure of restaurants, consumers have to cook at home by themselves, which also increases the demand for small kitchen appliances in China. This chapter is mainly Technology companies. Suggestions for improving the mature financial strategy are made. Broaden external financing channels in fundraising strategy, optimize the financing structure and reduce risks; do a good job in the budget, improve core competitiveness, and expand the industrial chain in mechanism in profit distribution to attract and retain core technical talents and improve financial performance evaluation. At the same time, for the implementation of the financial strategy, relevant investment strategy; establish a perfect incentive and distribution safeguards are proposed, including strengthening budget management;
REFERENCES

Design and Construction of Enterprise Financial Data Sharing Service Center Based on Big Data Technology

Yuxin Wang

wangy070322@163.com

Dongbei University of Finance and Economics, Dalian City, Liaoning Province, China

Abstract: With the rapid rise of digital economy, strengthening the construction and optimization of enterprise financial data sharing service center with the new generation of digital information technology is an important issue in the new period of high-quality development of enterprises. Therefore, this paper takes the financial data sharing service center as the research object, gives full play to the practical characteristics of big data technology, network information technology and computer application technology, and builds an enterprise financial data sharing service platform with Hadoop cluster as the data analysis and processing server and Javaweb as the interactive application system. The whole platform adopts B/S architecture design, and completes the development and deployment of various functional modules and API interfaces according to MVC pattern. On the basis of traditional accounting work, the platform will be further integrated into high-value financial work such as financial management analysis and cost review, which will make the platform's functional system more perfect and powerful, especially in the aspect of in-depth mining of financial data value. The system will give full play to its functional advantages and support K-means, CART decision tree and other data models to complete the identification and investigation of enterprise financial risks. The construction of the platform is not only conducive to promoting the realization of enterprise financial integration, but also makes outstanding contributions to improving the effectiveness of enterprise internal management.

Keywords: Big Data, Financial Data Sharing Service Center, Data Mining, Javaweb, Computer Application.

1 INTRODUCTION

Financial Shared Service Center (FSSC), as a brand-new enterprise financial operation and management mode, can highly meet the management requirements of multi-format or multi-branch group enterprises. The essence of the shared financial service center is financial centralization, that is, the accounting work of each branch is integrated, and the centralized operation is carried out in the service center, and all kinds of accounts and reports are automatically processed according to unified standards and norms, thus strengthening the financial control of the group headquarters over the branches. At the same time, it can also improve the efficiency of financial management, greatly reduce the financial system and labor costs, strengthen the sharing of data and information resources within enterprises, and promote the information transformation and upgrading of enterprise internal management.
However, the rapid popularization of digital technology and the rapid rise of digital economy have brought unprecedented influence to the economy and society. The new generation of digital technology represented by big data technology is permeating all industries and fields, making data a brand-new factor of production, and constantly promoting the emergence of new industries, new formats and new models, providing an important driving force for the transformation and upgrading of enterprises in the new era. [1] As a natural data center within an enterprise, the financial department can directly grasp all kinds of data and information resources, which is the key to the innovation and integration of big data technology. Similarly, the financial sharing service center should combine the characteristics and development trend of big data technology, promote the financial sharing service center to gradually evolve into a financial data sharing service center, and realize the update and improvement of the functional system of the financial sharing service center. It provides a set of feasible solutions for the actual effect, security and value problems in the process of data storage, call and mining. In view of this, this paper holds that in the new era, the enterprise financial sharing service center should adhere to innovation drive, give full play to the application advantages of big data technology, network information technology and computer application technology, and build an enterprise financial data sharing service platform with Hadoop cluster as data analysis and processing server and Javaweb as interactive application system. The whole platform is presented as a Web application, which can widely support enterprise managers, financial personnel and other business personnel to log in and use, remould various work processes, and overcome the shortcomings of inconvenient user operation and cumbersome business initiation process. It organically combines the financial system with the business system, strengthens the management of data interfaces among modules, promotes the circulation and interaction of various data, and realizes the sharing and utilization of resources, thus promoting the value creation of the integration of industry and finance under the financial sharing mechanism of enterprises, and making a beneficial attempt to improve the effectiveness of internal management of enterprises. [7]

2 INTRODUCTION OF KEY TECHNOLOGIES

2.1 Big Data Technology

With the rapid development and wide application of network information technology, Big data gradually emerges and rises, which is a new stage of information technology extension and development. Big data is a concept of continuous and dynamic change, which generally refers to a huge amount of data set that cannot be acquired, managed and processed by traditional information technology and software and hardware tools in a tolerable time. [10] The characteristics of big data, such as huge scale, various data types, fast processing speed, low value density and complex content, determine the application direction and development trend of big data, and also put forward new requirements for data processing technology.

Big data technology, namely big data processing application technology, is a collection of technologies covering the whole life cycle of big data, and it is also a new mode different from traditional data processing. Big data technology includes collection, cleaning, storage, analysis and mining, visual display and application, which provides a complete processing paradigm for
the acquisition and reflection of the value of big data, and constantly expands the application scenarios of big data.

2.2 Hadoop Technical Framework

Hadoop is an open source technical framework implemented by Java language. Hadoop can run on large-scale clusters, can be used as the core of the underlying storage and analysis, and provide users with reliable, scalable and distributed computing big data services. \(^5\) The technical core of Hadoop is distributed storage and distributed computing, which are implemented by HDFS and MapReduce respectively.

2.2.1 Kafka

Kafka is in the transport layer of big data technology stack heap, which is essentially a distributed information queue. Its core application is to realize the asynchronous transmission and transmission of data and reduce the coupling between data sources and data storage units. \(^3\) Kafka has excellent data throughput, and can accept many clients or data sources at the same time. It is widely used in large-scale real-time data calculation and log collection and analysis application scenarios. Kafka architecture is mainly composed of producers, consumers and brokers.

2.2.2 Sqoop

Sqoop is a data migration component under Hadoop framework, and its target object is the conventional relational database, which aims at assisting the rapid and direct data transmission and transfer between RDBMS and Hadoop. Sqoop can support the bidirectional flow of structured data, that is, data can be imported into HDFS under Hadoop from Mysql, Oracle, Postgresql and other databases, or data can be extracted from Hadoop framework and exported to the corresponding database. The underlying program of Sqoop is extracted, transformed and loaded by MapReduce, which can ensure that Sqoop has high concurrency efficiency and high fault tolerance in the running process, and greatly improve the performance of Hadoop in calling data resources. \(^9\)

2.3 JavaWeb

JavaWeb is the sum total of technologies that use Java technology to solve related Web and Internet fields, and it is the core technology that uses Java language to realize dynamic Web application development. The key work of JavaWeb application development lies in the design and development of background Web Server, that is, the construction and deployment of each functional module, business logic control and database application model. The development of JavaWeb has gone through many stages in different periods, from the earliest Servlet technology to the subsequent JSP technology, and then to the extensive application of a large number of development frameworks. The development of Web applications is further hierarchical, structured and modular, so as to make the development of Web applications more efficient and concise. \(^6\)

2.4 Data Mining

Data mining (DM), as a kind of computer science and technology, aims at the complex process of extracting and mining the hidden and valuable patterns or laws from a large number of
incomplete and noisy actual data. The birth of data mining technology meets people's demand for the analysis and processing of massive information data in the current information explosion era. It is a technology that integrates and comprehensively applies database technology, information science, statistics, machine learning, data timeliness and other technologies. Compared with traditional data analysis, data mining technology can handle more complex data objects, including structured data, semi-structured data and heterogeneous data. The main process is shown in Figure 1, in which database, data warehouse, Internet and other data sources represent the collection of raw data, while the data preprocessing stage includes three basic steps: data cleaning, data integration, data selection and transformation.

![Figure 1: Data mining process](image)

The construction of data mining model is the core of the whole data mining work. The construction of data mining model corresponds to data analysis method, which not only determines the application direction of data mining results, but also determines the construction of data mining model. Common data analysis methods include classification analysis, prediction analysis, cluster analysis, valuation analysis and correlation analysis, as shown in Table 1.

<table>
<thead>
<tr>
<th>Classification</th>
<th>Data analysis method</th>
<th>Data mining model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guided data mining</td>
<td>Classification analysis</td>
<td>Decision tree, random forest, neural network</td>
</tr>
<tr>
<td></td>
<td>Prediction analysis</td>
<td>Regression tree and rough set method</td>
</tr>
<tr>
<td></td>
<td>Valuation analysis</td>
<td>SVM, bayesian method</td>
</tr>
<tr>
<td>No guided data mining</td>
<td>Correlation analysis</td>
<td>Pearson correlation coefficient and aprior algorithm</td>
</tr>
<tr>
<td></td>
<td>Cluster analysis</td>
<td>K-means clustering algorithm, lineage clustering</td>
</tr>
</tbody>
</table>

### 2.5 Development Process

According to the application requirements of the above related application technologies, complete the configuration and deployment of the development environment of the enterprise financial data sharing service platform. The development content of the system is divided into two parts. One is to build Hadoop cluster to complete the collection and storage management of financial data. Secondly, under the Java development environment, the development of Web Server is completed with Spring framework, and the business system, financial system and data analysis system of the financial sharing service center are built to form a standard Web application.
First of all, Hadoop cluster architecture needs the support of hardware and software. The underlying operating system is Linux, CentOS 6.7(x86_64) is the version, and jdk-8u291-linux-x64 is the JDK version. According to the application requirements of the system, Hadoop cluster will be set up into seven nodes, named Master1, Master2, Slave1, Slave2, Slave3, Slave4 and Slave5 respectively. Master is the master node and Slave is the slave node. Hadoop version is 2.7.7, which is installed in each node, and components such as Yarn, HDFS, Zookeeper, HBase and Kafka are also deployed in each node. In addition, the system will use Sqoop component to import financial data, business data and other data into HDFS system under Hadoop framework to realize distributed storage and form original data. With the help of Java, the sample information table is constructed, and the data preprocessing operation is completed and imported into MySQL to obtain sample data.

Secondly, for the development of each functional module of the shared service platform, the workflow engine will be introduced to complete the process service control and performance expansion of each function. The system selects Activiti open source workflow engine, and gives the core class ProcessEngine process engine and some Service classes in Activiti to Spring container for management, so as to realize the call and processing of workflow by Web Server. In addition, for the data mining analysis function, the system will build different data algorithm models and package them into classes that can be called directly, as shown in Figure 2, which is the key code to realize the K-means clustering algorithm model.

```
public class KMeans
{
    public List<Point> points;
    FileWriter out = null;
    DecimalFormat dfFormat = new DecimalFormat("00.00");
    public KMeansCluster KMeansCluster;
    public int numClusters = 5;
    public int numPoints = 50;
    public static final String FILEPATH = "F:\kmeans\test.txt";
    public static void main(String[] args)
    {
        LineInitList = new LineList(100, 2, 200);
        KMeans InitKMeans;
        KMeans.printRes();
        KMeans.saveResToFile(FILEPATH);
    }
}
```

**Figure 2:** Model building code of the K-means clustering algorithm

Finally, for the development of Web application server, the basic development environment is Java, JDK version requires 1.6 or above, Java development environment is MyEclipse 2014, Web server is Tomcat 8.0, and database server is MySQL 5.5. And the project object model (Maven) is used to manage the project structure. Maven chooses Apache-Maven-3.2.1 version. In the process of building the overall development environment, the installation of JDK and the configuration of environment variables are completed first to build the foundation of Java application development. Secondly, the installation of MyEclipse and the installation of Tomcat, the Web server, and the configuration of Tomcat is completed in the Preference option under MyEclipse. Then, based on Spring architecture, the integration and encapsulation of the whole system is completed. Through the introduction of the above key technical theories, we have determined the overall environment of system development, the configuration of related
software and tools, and the technical feasibility of the overall project of enterprise financial data sharing service platform.

3 FUNCTIONAL IMPLEMENTATION

3.1 Business System

Under this function module, the platform will plan and integrate all businesses such as procurement, asset management, customer relationship management, project management, logistics and warehousing, sales and operation according to the standardized process formulated by the enterprise, and design corresponding forms according to different business properties, each of which contains several key fields. When the user initiates the business process, the user can input, upload or scan the corresponding content according to the form. The system automatically starts the workflow engine according to the standardized business process to accept the business data and forward it to the approval verification node. After the approval verification is correct, the business data and the approval result will be returned to the business system for display. At the same time, the approval and verification node converts the business form into financial information voucher, and transmits it to the financial system. After review, the receipt and payment of funds are completed.

3.2 Financial System

Under this functional module, the platform will be highly compatible with the settings of various functional modules of the enterprise financial management software, mainly providing functional interfaces such as voucher management, account table inquiry, summary accounting and report filling. After financial personnel log in to the system, they can intuitively see the initiation and execution of various business processes, and the corresponding vouchers and forms will be naturalized into different subsystems according to their categories, which will become the data source of the financial data sharing service platform, thus realizing data sharing and providing the foundation for subsequent data value mining and application.

3.3 Data Analysis System

Under this function module, system users can synthesize financial data, calculate financial indicators and evaluate risk control for related projects, data information and departments. For example, after selecting different departments, projects, time and other basic information in the system in turn, users can calculate the analysis indicators from four aspects: business ability, debt paying ability, profitability and equity analysis according to the analysis and calculation rules. Table 2 shows some index information, in which KMO is the measured value, which provides the necessary weight proportion for the subsequent calculation of scores. After calculation, the enterprise's abilities are graded accordingly, and this is the key feature, and the K-means clustering algorithm model is selected to complete the data classification, thus generating the classification standard, which is compared according to the standard. If the distribution of a certain ability is quite different from that of previous years, the existence of financial risks can be determined. As shown in Table 3, the analysis table of enterprise financial risk clustering results shows that the greater the positive index value, the greater the financial risk, and the greater the reverse index value, the smaller the financial risk.
Table 2: Calculation formulas of some indexes

<table>
<thead>
<tr>
<th>Category</th>
<th>Specific items</th>
<th>Analysis index</th>
<th>Computational formula</th>
<th>KMO factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operation capacity</td>
<td>Assets</td>
<td>Proportion of current assets</td>
<td>Current assets / Total assets</td>
<td>0.188</td>
</tr>
<tr>
<td></td>
<td>Cash flow</td>
<td>Cash turnover rate</td>
<td>Main business income / Cash balance</td>
<td>0.291</td>
</tr>
<tr>
<td>Debt paying ability</td>
<td>Creditor's rights</td>
<td>Quick ratio</td>
<td>(current assets-liabilities)/Current liabilities</td>
<td>0.271</td>
</tr>
<tr>
<td>Profitability</td>
<td>Profit contributio n</td>
<td>Net interest rate on sales</td>
<td>Net profit / Total profit</td>
<td>0.213</td>
</tr>
<tr>
<td>Equity analysis</td>
<td>Owner's equity</td>
<td>Total asset growth rate</td>
<td>Total assets year-on-year</td>
<td>0.403</td>
</tr>
</tbody>
</table>

Table 3: Analysis Table of Enterprise Financial Risk Clustering Results

<table>
<thead>
<tr>
<th>Category</th>
<th>Analysis index</th>
<th>Cluster 1</th>
<th>Cluster 2</th>
<th>Cluster 3</th>
<th>Evaluate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operating ability</td>
<td>Current assets ratio</td>
<td>3.38</td>
<td>2.11</td>
<td>1.56</td>
<td>Reverse index</td>
</tr>
<tr>
<td></td>
<td>Cash turnover ratio</td>
<td>11.84</td>
<td>4.66</td>
<td>7.85</td>
<td>Reverse index</td>
</tr>
<tr>
<td>Solvency</td>
<td>Quick ratio</td>
<td>0.15</td>
<td>0.33</td>
<td>0.17</td>
<td>Reverse index</td>
</tr>
<tr>
<td>Profitability</td>
<td>Net profit rate of sales</td>
<td>0.21</td>
<td>0.32</td>
<td>0.22</td>
<td>Reverse index</td>
</tr>
<tr>
<td>Equity analysis</td>
<td>Total asset growth rate</td>
<td>0.66</td>
<td>0.61</td>
<td>0.26</td>
<td>Positive index</td>
</tr>
</tbody>
</table>

4 CONCLUSIONS

With the purpose of promoting the transformation of the operation mode of enterprise shared financial service center, this paper constructs a Web-based enterprise financial data sharing service platform based on the functional characteristics of big data technology, network information technology and computer application technology. The function of the platform combines the business system and financial system of an enterprise, completes the sharing and interaction of all kinds of data, and provides a brand-new data analysis system. It is convenient for users to improve the processing efficiency of various businesses with concise, convenient and efficient operation, increase the value embodiment and rational application of financial data, strengthen the internal financial management ability of group enterprises, and promote the digitalization and intelligent construction of internal management and control of enterprises.
REFERENCES

Research on the Influence of Digital Capability on Service Transformation of Manufacturing Enterprises*

Hailin Dong¹,a, Juhong Chen¹

Email: donghailin123@163.com, Tel:18189503628¹

School of Economics and Management, Xi’an University of Technology, Xi’an 710054, China¹

Abstract: Digital capability is an important support for service transformation of manufacturing enterprises. But there is a lack of in-depth research on how to use digital capability to promote service transformation. Based on the dynamic capability theory, through the questionnaire survey of 215 manufacturing enterprises, this paper explores the internal mechanism of digital capability driving the service transformation of manufacturing enterprises by using the analytic hierarchy process and Bootstrap method. The research shows that:(1) digital capability promotes the service transformation of manufacturing enterprises, and resource flexibility and coordination flexibility play the full mediation effect;(2) More interestingly, of the effect size of different mediation paths, coordination flexibility has the strongest effect. The research conclusion not only presents a new theoretical explanation for the mixed findings on the relationship between digital capabilities and firm performance, but also provides a useful reference for the service transformation of manufacturing enterprises.

Keywords: Digital Capability, Resources Flexibility, Coordination Flexibility, Service Transformation.

1 INTRODUCTION

In recent years, with the development of digital technologies such as Internet of Things, big data and cloud computing, they have brought revolutionary power source to the service of manufacturing enterprises [¹]. Driven by digital technology, the service business of manufacturing enterprises has been transformed from traditional experience-based passive response service to data-driven accurate prediction and intelligent decision-making service [⁸]. Digital capability is an important support for the service transformation of manufacturing enterprises in the era of digital economy [¹¹]. By improving digital capability, enterprises constantly optimize processes, improve organizational agility [¹²], innovation capability and optimize human capital structure [²⁴] to empower enterprise value creation. However, Kohtamäki (2020) believe that digital capability does not always improve organizational performance. Based on this, in digital economy, it has become an important research topic to clarify the action mechanism of digital capability on service transformation and promote the high-quality development of manufacturing enterprises.

* This paper is supported by The National Social Science Fund of China:22BTJ050.
At present, the academic research on service-oriented manufacturing enterprises mainly focuses on organizational change and resource allocation [2], service innovation capability [21], service operation and maintenance business model innovation and other issues [3]. In recent years, with the increasing maturity of digital technology, the influence of digital technology on service has attracted the attention of some scholars. However, there is still a lack of research on how to use digital capabilities to drive the service transformation of manufacturing enterprises. Thus, this study takes 215 manufacturing enterprises that are using digital technology for service transformation as samples. Based on the dynamic capability theory, this study tries to open the "black box" between digital capability and service transformation performance of manufacturing enterprises from the perspective of strategic flexibility by using the analytic hierarchy process and Bootstrap method.

This study contributes to the existing literature in the following two aspects: First, based on the dynamic capability view, we explore the internal mechanism of digital capability driving the service transformation of manufacturing enterprises, and further enriching the related research of manufacturing service; The existing literature emphasizes the promotion of digitalization to enterprises' slight flexibility, but the research regards strategic flexibility as a whole, and does not distinguish the difference between resource flexibility and coordination flexibility. Through empirical analysis, this paper holds that coordination flexibility plays a more significant mediating role between digitalization capability and service innovation performance than resource flexibility. This discovery provides new guidance for the service transformation of manufacturing enterprises.

2 THEORY AND HYPOTHESES

2.1 Dynamic Capability View

Dynamic capability refers to the ability of perceiving and seizing new opportunities, reconfigure and protect knowledge assets, capabilities and complementary assets, so as to achieve sustainable competitive advantage. The organizations try to adapt to the new environment by identifying and capturing new strategic opportunities, coordinating necessary organizational assets, inventing new business models and organizational forms, changing business processes and restructuring resources [19]. In the digital economy, the reconstruction and integration of internal and external resources are promoted by embedding digital technology into organizations and processes, so as to promote the innovation of processes [9]. Digital capability reflects the perception dimension of dynamic capability to a certain extent. Resource flexibility and coordination flexibility are conducive to enterprises to seize the opportunity to reconfigure resources to cope with the uncertainty of the external environment.

2.2 Digital Capability and Service Transformation Performance

Digital capability refers to the ability of enterprises to reconstruct the existing resources, structure, value, boundaries and other factors through digital technology, and cultivate and construct the ability to transform the production factors of digital economy into business model innovation in the process of digital transformation [17]. It mainly includes digital perception capability, digital operation capability and digital collaborative operation capability [22]. LI et al. (2022) think that enterprises with digital capabilities can strengthen the dynamic scanning of
the external environment through new digital channels (e.g., digital platforms, digital systems, etc.), perceive the changes of external environment and consumer demand in time, seize opportunities in time through market capitalization agility and operational adjustment agility, and reconfigure existing resources and processes to improve enterprise performance. Lenka et al. (2017) deconstruct the digital capability into three dimensions: digital intelligence, connection and analysis. We can give full play to the role of customer value supply to realize enterprise value co-creation through the perception and response the customers. Hence, we propose the following:

Hypothesis 1. Digital capability positively affects the performance of service transformation.

2.3 Digital Capability and Strategic Flexibility

Strategic flexibility refers to the ability of an enterprise to dynamically identify and respond to changes in the external environment in time to reshape organizational resources and strategies to improve its competitive advantage [14], including resource flexibility and coordination flexibility [18]. In the digital economy, more and more enterprises are speeding up the rapid integration of existing resources and business processes by cultivating their own digital ability, and improving the ability of enterprise resource integration. On the other hand, driven by the new generation of information technology, enterprises are more conducive to breaking the traditional organizational boundaries, and improving the circulation efficiency of resource elements and the flexibility of resource allocation through cross-border search in the R&D process; On the other hand, enterprises use artificial intelligence algorithms, mobile Internet and other digital technologies to explore new business channels, use the value network space to find key resources, and establish a flexible organizational structure, thus improving the flexibility of organizational coordination [15]; Some scholars also believe that in the turbulent market environment, the improvement of digital capability is conducive to the innovation of digital business model, which in turn affects the flexibility of enterprise strategy matching with the innovation of business model, thus realizing sustainable value acquisition and value creation [4]. Accordingly, we postulate the following:

Hypothesis 2a. Digital capability positively affects resource flexibility.

Hypothesis 2b. Digital capability positively affects coordination flexibility.

2.4 Mediating Role of Strategic Flexibility

The dynamic capability view suggests enterprises with digital capability can dynamically perceive the changes of external environment, and improve their responsiveness to external environment by coordinating and integrating internal resources, processes and other elements. For example, Enrique et al.(2022) believe that enterprises with digital capabilities can use big data, Internet of Things and other technologies to improve the flexibility of internal supply chain operation, realize the operational efficiency of products from production to delivery, and improve operational performance. Toorajipour et al.(2021) think that enterprises can use machine learning, big data, cloud computing and other methods to strengthen the dynamic monitoring of the whole product production process, predict the future customer behavior, shorten the delivery time and reduce logistics costs. At the same time, the use of artificial intelligence technology makes the coordination between suppliers and distributors more flexible. However, only relying on digital capabilities to obtain external data can't directly help
companies improve their performance, the organizations need to be able to flexibly identify opportunities and adjust and restructure existing business processes, so as to promote the growth of enterprise performance \cite{12}. Strategic flexibility is just such an ability to adapt to the changes of market demand through the flexible allocation of resources, the adjustment of organizational elements and processes. Hence, we pose the following:

**Hypothesis 3** Strategic flexibility mediates the relationship between digitalization ability and service transformation performance.

### 3 METHODS

#### 3.1 Data Collection

Before the formal distribution of the questionnaire, this study learned through a series of interviews that the managers of manufacturing enterprises such as digital department, information management, data application and R&D department can become potential research objects of this study, and they can deeply participate in big data management and service innovation. A total of 435 questionnaires were distributed. After eliminating invalid questionnaires such as similarity, obvious errors and short filling time, 215 valid questionnaires were finally obtained, with an effective recovery rate of 49.4%. The samples were mainly from Guangdong, Zhejiang, Shanxi, Fujian, Jiangsu and other places.

<table>
<thead>
<tr>
<th>Item</th>
<th>item description</th>
<th>sample</th>
<th>percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Establish years</td>
<td>5years</td>
<td>49</td>
<td>22.70%</td>
</tr>
<tr>
<td></td>
<td>6-10 years</td>
<td>63</td>
<td>29.08%</td>
</tr>
<tr>
<td></td>
<td>11-15 years</td>
<td>44</td>
<td>20.57%</td>
</tr>
<tr>
<td></td>
<td>&gt;16 years</td>
<td>59</td>
<td>27.66%</td>
</tr>
<tr>
<td>Scale</td>
<td>100 ≤ Size&lt;500</td>
<td>65</td>
<td>30.26%</td>
</tr>
<tr>
<td></td>
<td>500 ≤ Size&lt;1500</td>
<td>78</td>
<td>36.41%</td>
</tr>
<tr>
<td></td>
<td>1500 ≤ Size</td>
<td>72</td>
<td>33.33%</td>
</tr>
<tr>
<td>Ownerships</td>
<td>State-owned</td>
<td>49</td>
<td>22.70%</td>
</tr>
<tr>
<td></td>
<td>Privately owned</td>
<td>97</td>
<td>44.92%</td>
</tr>
<tr>
<td></td>
<td>Foreign</td>
<td>26</td>
<td>12.06%</td>
</tr>
<tr>
<td></td>
<td>Others</td>
<td>43</td>
<td>20.33%</td>
</tr>
</tbody>
</table>

#### 3.2 Measures

In order to ensure the validity of the measurement, all variables adopt the mature scale, which is revised and formed into the initial scale after translation and back translation. Through a small sample test of EMBA and MBA students in our school, some expressions in the initial scale are revised and adjusted according to the collected 52 questionnaires and the relevant feedback of the respondents, and finally the measurement scale is determined. Each variable item is evaluated by Likert5 scale. Among them, the digital capability reference to the scales of Lenka

3.3 Reliability and Validity

In terms of reliability, Cronbach's Alpha coefficients of digitalization ability, resource flexibility, coordination flexibility and service transformation performance are 0.921, 0.853, 0.889 and 0.910, and the combined reliability CR is 0.923, 0.854, 0.890 and 0.876, which are all higher than the good threshold level of reliability test 0.8. In terms of validity, the factor load of each item exceeds 0.7, and the mean variance extraction (AVE) is greater than 0.65, which indicates that the questionnaire has high convergence validity. We test the structural validity of the model and the distinguishing validity between variables by selecting $\chi^2$/df, RMSEA, SRMR, CFI and TLI indicators. Table 2 shows that the questionnaire has good discrimination validity.

<table>
<thead>
<tr>
<th>Model</th>
<th>Factor</th>
<th>$\chi^2$/df</th>
<th>RMSEA</th>
<th>SRMR</th>
<th>CFI</th>
<th>TLI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Four-factor model</td>
<td>DC, RF, CF, STP</td>
<td>1.29</td>
<td>0.03</td>
<td>0.03</td>
<td>0.98</td>
<td>0.98</td>
</tr>
<tr>
<td>Three-factor model</td>
<td>DC, RF+CF, STP</td>
<td>2.91</td>
<td>0.07</td>
<td>0.05</td>
<td>0.89</td>
<td>0.88</td>
</tr>
<tr>
<td>Two-factor model</td>
<td>DC, RF+CF+STP</td>
<td>3.90</td>
<td>0.08</td>
<td>0.07</td>
<td>0.83</td>
<td>0.81</td>
</tr>
<tr>
<td>Single-factor model</td>
<td>DC+RF+CF+STP</td>
<td>6.46</td>
<td>0.11</td>
<td>0.10</td>
<td>0.68</td>
<td>0.65</td>
</tr>
</tbody>
</table>

Note: DC, RF, CF, STP respectively represent digitalization capability, resource flexibility, coordination flexibility and service transformation performance. “+” means factor model merging.

3.4 Common Method Bias

The statistical method of Harman single factor test is adopted to test, and the variance interpretation rate of the largest factor is 21.791%, which is less than 40%, indicating that there is no obvious common method deviation problem.

3.5 Hypothesis Testing

From Table 3, it can be seen that first, digital capability has a significant positive impact on resource flexibility ($\beta=0.434$, $P<0.001$), and coordination flexibility ($\beta=0.530$, $P<0.001$). The results support Hypotheses 2a and 2b. Second, digital capability ($\beta=0.411$, $P<0.001$) has a significant positive impact on service transformation performance when there are no presumed mediators. Model 4 is the whole model after adding the intermediary variable of strategic flexibility, and only the regression coefficient of resource flexibility ($\beta=0.292$, $P<0.05$) and coordination flexibility ($\beta=0.335$, $P<0.01$) is significant, however, such a positive association becomes non-significant ($\beta=0.203$, $P>0.10$) when the presumed mediators exist. The above results thus confirm the full mediation effect of resource flexibility and coordination flexibility. To ensure robustness, Table 4 reveals the results derived from the bootstrap method. Because the 95% confidence interval of the direct path includes zero, whereas the 95% confidence intervals of the two indirect paths exclude zero, these results confirm the full mediation effect of resource flexibility and coordination flexibility again. Overall, our results support Hypotheses 3. At the same time, the path difference between coordination flexibility and resource flexibility is 0.1029, and the confidence interval is [0.692, 0.1815], excluding 0, which indicates that
coordination flexibility has a greater impact on different intermediary paths.

### Table 3 Regression model and results

<table>
<thead>
<tr>
<th>Variable</th>
<th>RF</th>
<th>CF</th>
<th>STP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
<td>Model 3</td>
</tr>
<tr>
<td>Establish years</td>
<td>-0.005</td>
<td>-0.086</td>
<td>0.016</td>
</tr>
<tr>
<td>Size</td>
<td>0.041</td>
<td>0.077</td>
<td>0.005</td>
</tr>
<tr>
<td>Ownerships</td>
<td>-0.056</td>
<td>0.015</td>
<td>-0.012</td>
</tr>
<tr>
<td>Industry</td>
<td>0.007</td>
<td>-0.005</td>
<td>-0.017</td>
</tr>
<tr>
<td>DC</td>
<td>0.434***</td>
<td>0.530***</td>
<td>0.411***</td>
</tr>
<tr>
<td>RF</td>
<td>0.292*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CF</td>
<td></td>
<td></td>
<td>0.292*</td>
</tr>
<tr>
<td>R2</td>
<td></td>
<td></td>
<td>0.435**</td>
</tr>
<tr>
<td>F</td>
<td>21.95***</td>
<td>29.75***</td>
<td>27.26***</td>
</tr>
</tbody>
</table>

Note: ***, **, and * represent significance at 0.001, 0.01, and 0.05 level, respectively.

### Table 4 The effects of digital capability on service transformation innovation.

<table>
<thead>
<tr>
<th>paths</th>
<th>Effect</th>
<th>BootSE</th>
<th>BootLLCI</th>
<th>BootULCI</th>
</tr>
</thead>
<tbody>
<tr>
<td>DC→STP</td>
<td>0.073</td>
<td>0.046</td>
<td>-0.018</td>
<td>0.165</td>
</tr>
<tr>
<td>DC→RF→STP</td>
<td>0.1022</td>
<td>0.037</td>
<td>0.1196</td>
<td>0.2642</td>
</tr>
<tr>
<td>DC→CF→STP</td>
<td>0.578</td>
<td>0.034</td>
<td>0.0511</td>
<td>0.1234</td>
</tr>
</tbody>
</table>

4 CONCLUSION

Based on the sample of 215 manufacturing enterprises that are using digital technology for service transformation, this paper analyzes the mechanism of strategic flexibility between digital capability and service transformation of manufacturing enterprises by using the analytic hierarchy process and Bootstrap method. The results show that: (1) Digital capability promotes the service transformation of manufacturing enterprises, and resource flexibility and coordination flexibility play the full mediation effect; (2) More interestingly, of the effect size of different mediation paths, coordination flexibility has the stronger effect.

5 MANAGERIAL IMPLICATIONS

Based on the above conclusions, this paper puts forward the following suggestions: First, manufacturing enterprises should seize the opportunity of digital transformation and vigorously carry out digital capacity building. Accelerate the innovation and application of big data, artificial intelligence, and the mobile Internet in R&D, production, marketing and other links, and improve the digital operation capability. Second, give full play to the transmission function of resource flexibility and coordination flexibility between digitalization capability and service transformation. Build a big data platform or network, strengthen communication and
collaboration with stakeholders to form an innovation ecosystem, shorten the development process of new products or services, and thus improve service innovation performance.

REFERENCES


Research on Synthetic Evaluation Model for Enterprise Integration Development

Yuejia Sun\textsuperscript{1, a*}, Hongyun Pan\textsuperscript{1, b}, Weixu Liu\textsuperscript{1, c}, Zhaoxi Wang\textsuperscript{1, d}, Lin Lin\textsuperscript{1, e}

*sunyuejia@chinamobile.com, panhongyun@chinamobile.com, liuweixu@chinamobile.com, wangzhaoxi@chinamobile.com, linlinyf@chinamobile.com

Department of User and Market Research China Mobile Research Institute Beijing, China

Abstract: With the development of digital economy, multi-business integration has become the development strategy of many enterprises. Multi-business integration of telecom enterprises refers to the provision of telecom products and services covering multiple fields of CHBG to customers in a "package" form. Under this development trend, the synthetic evaluation of integrated development level has become a very meaningful topic in enterprise management. In this paper, an evaluation model of subjective and objective fusion for enterprise integration development is constructed. Firstly, an AOE evaluation factor system is proposed based on the development requirements. Then, the first level index evaluation method based on AHP is designed, and the second level index evaluation method based on entropy weight method, CRITIC method and fuzzy synthetic evaluation method is designed. Finally, a synthetic evaluation model of subjective and objective fusion is constructed to fully integrate the experience of experts and the information of the statistical data. In terms of application, through the investigation data and sampling data of some provincial companies of China mobile, the model was applied to obtain the evaluation results of the enterprise in three aspects: integration capability, integration operation and integration effectiveness. The empirical analysis results show that integration development can promote customer value and loyalty. Full integration customers (C+H+B) have the highest value increase rate of 92% and the highest loyalty increase rate of 68%. In conclusion, the evaluation model can effectively assist the scientific decision-making in the management and operation process of enterprise integration development.

Keywords: Multi-Business Integration, Synthetic Evaluation Model, Subjective and Objective Fusion, AHP, Entropy Weight Method, CRITIC, Fuzzy Synthetic Evaluation.

1 INTRODUCTION

With the development of the digital economy, the integration of multi-business has become the development strategy of many enterprises, which can promote the efficient allocation of new elements such as technology and data, and realize the aggregation and sharing of resources and elements\textsuperscript{1,2}. Multi-business integration for telecom enterprises refers to the provision of telecom products and services covering multiple areas of CHBG to customers in the form of a "package"
combination, which is a business model that implements sales, billing, and service by product portfolio.

Under this development trend, the synthetic evaluation of integration development level has become an important topic. Evaluation is an important prerequisite for scientific decision. Therefore, it is necessary to construct an evaluation model for the level of enterprise integration development, so as to achieve the goals of goal guidance, quantitative evaluation and comprehensive scoring. First, achieve the goal of direction guidance. By dismantling and quantifying the integration development goals, it points out the direction for the company's development. Second, realize quantitative assessment. Through the quantitative evaluation of the current development status of integration, the current development progress can be presented. Third, calculate the overall scores. Based on the comprehensive scoring results, the subsidiaries are compared horizontally to help them identify the problems encountered in the integration development.

Scholars at home and abroad have carried out a series of studies on comprehensive evaluation. In the 70s and 80s of the 20th century, a variety of widely used evaluation methods appeared, such as analytic hierarchy method [3], data envelopment analysis method [4], and TOPSIS method. In the 80s of the 20th century, many new evaluation ideas and theories of gray system theory [5], artificial neural network technology [6], and information theory were continuously integrated with traditional methods, and new evaluation methods and models were born. After the 90s of the 20th century, the combination of evaluation methods has become a hot spot in the field of evaluation, and methods such as fuzzy artificial neural network system, AHP-entropy [7], AHP-fuzzy evaluation method [8] have been produced, which have enriched the research results of synthetic evaluation methods. In terms of enterprise evaluation studies, Yufang Sun proposed a synthetic evaluation method based on analytic hierarchy and fuzzy evaluation method for the telecom operators' services [9], Xiaotong Pang used entropy-weight TOPSIS to evaluate the triple-network fusion [10]. Li Anmin researched the operation mode and evaluation system of Chinese telecom operators, and constructed a model of the impact factors of technology, market and value creation system [11].

Based on the above research, this paper constructs a subjective and objective fusion evaluation model for the development of enterprise integration, which has the advantage of fully integrating expert experience and data measurement results. Finally, the model application and empirical analysis are carried out through the survey and sampling data of some provincial companies, which verifies the scientificity and feasibility of the model.

2 EVALUATION FACTOR SYSTEM CONSTRUCTION

2.1 Principles for the Construction of Factor System

Based on the requirements of enterprise integration development planning, an evaluation factor system is designed. The factor system is the basis for reflecting the level of integration development, which should meet the principles of consistency, adequacy, systematicness, feasibility, comparability, dynamism and predictability.
• Consistency: Ensure that indicators reflect the strategic direction of integration development and are consistent with the planning goals.

• Adequacy: Multiple perspectives of process and results should be included, covering the construction process, operation level, implementation effect, etc.

• Systematic: There is a hierarchical logical relationship between the indicators, from top to bottom, from macro to micro. The systematic structure that supports structured dismantling can assist in the problem location in the process of integration development.

• Feasibility: Indicators in the evaluation system should be clearly defined, measurable and make full use of available statistics.

• Comparability: Design relative values and uniform measurement schemes to achieve horizontal and vertical comparisons of different metrics.

• Dynamic: Establish a dynamic adjustment mechanism for indicators to meet market changes.

• Predictability: Choose indicators with long-term effectiveness and business foresight to meet forecasting and decision-making needs.

2.2 Influencing Factor Model

Following the above factor system principles, a factor model consisting of three dimensions: integration capability evaluation, integration operation evaluation and integration effect evaluation is constructed, namely the AOE model.

![AOE model](image)

**Figure 1.** AOE model for influencing factor system

The ability dimension reflects the progress of enterprises in building integrated products, organizations, contact points, platforms, resources, and assessment. The operation dimension reflects the operation status of typical converged business scenarios of enterprises. The effect dimension reflects the development effect in terms of customer scale, value, and loyalty.
The evaluation indicator system should have a dynamic iterative mechanism, including indicator entry, indicator exit, and indicator replacement.

- **Indicator Entry Mechanism**
  - Ability evaluation indicators: In the actual operation process, new indicators are supplemented according to the new needs of pre-sales, sales, and after-sales processes.
  - Operation evaluation indicators: For newly added fusion scenarios, data statistics are performed according to the standard scenario indicator template, and the new scenario indicators are included in the fourth-level structure of the indicator tree.

- **Indicator Exit Mechanism**
  - Ability evaluation indicators: If an indicator reaches 100% in all provinces and the indicator area issues its differentiation, such indicators can be deleted.
  - Operation evaluation indicators: For relatively mature integration scenarios, if the completion rate of each province reaches a certain threshold, it can be deleted from the evaluation indicators.

- **Indicator Replacement Mechanism**
  - Ability evaluation indicators: According to the needs of different stages of integration development, gradually replace statistical indicators that meet the current degree of integration.
  - Operation evaluation indicators: Choose from a combination of metrics such as Cumulative Number, New Additions, Net Increase, Active Count, Target Completion Rate, Overall Coverage, Market Share, etc. For example, in the early stage of business development, focus on indicators such as cumulative number, new additions, active numbers, and retention. During the mature period of business development, focus on indicators such as activity effectiveness and market share.
3 SYNTHETIC EVALUATION MODEL

The evaluation model is constructed by subjective and objective fusion methods, and the expert experience and data measurement results are fully integrated. Based on subjective experience driven, the model can effectively meet the needs of converged development goals. At the same time, based on the data measurement results, the scientificity of the model can be improved.

3.1 Subjective Evaluation Method

AHP analytic hierarchy is a decision-making weight research method that combines qualitative and quantitative to solve multi-factor complex problems. Based on the experience of decision makers, this method determines the relative importance of each measurement element, and calculates the weight of each element based on statistical methods.

Step 1: Scale determination and judgment matrix construction. In this scheme, the 1-5 degree scale method is used to obtain the judgment matrix $A$ through expert scoring.

Step 2: Feature vectors and weights calculation. The columns of the judgment matrix are summed and normalized to obtain the $B$ matrix. The rows are summed to obtain the feature vector $C$ matrix, and then the $C$ matrix is normalized to obtain the weights.

$$b_{ij} = a_{ij} / \sum_i a_{ij}, \ c_i = \sum_j b_{ij}$$ (1)

Step 3: Consistency test analysis. The consistency analysis is carried out by calculating the maximum feature root, $CI$ value, $RI$ value, and $CR$ value to avoid the logic error of the judgment matrix. If the $CR$ value is less than 0.1, it means that the consistency test has passed.

Step 4: Analysis and weight fine-tuning. After the consistency test passes, the final applied weight result is determined by fine-tuned processing.

$$\lambda_{\text{max}} = \sum_i \frac{a_i W_i}{\sum_i a_i W_i}, \ CI = \frac{\lambda_{\text{max}} - n}{n-1}, \ CR = \frac{CI}{RI}$$ (2)

3.2 Objective Evaluation - Entropy Weight Method

Entropy is a physical unit of measurement. Higher entropy indicates more chaotic data, less information carried, smaller utility values, and therefore smaller weights. The entropy method is a research method that combines the information value provided by the entropy value to determine the weight, avoiding the bias caused by human factors. Compared with the subjective assignment method, the entropy method has higher accuracy and stronger objectivity, which can better explain the results obtained.

Step 1: Indicator forwardization and data standardization. Convert all indicators into positive indicators and carry out forward processing of indicators. Then the data standardization process is used to balance the errors caused by the differences between indexes or dimensionality.

$$Z_{ij} = \frac{x_{ij} - \min(x_{1j}, x_{2j}, \ldots, x_{nj})}{\max(x_{1j}, x_{2j}, \ldots, x_{nj}) - \min(x_{1j}, x_{2j}, \ldots, x_{nj})}$$ (3)
Step 2: Indicator information entropy calculation. Calculate the probability matrix $p$ of the indicators, and then calculate the information entropy $e$ of each indicator.

$$p_{ij} = \frac{z_{ij}}{\sum_{i=1}^{n} z_{ij}}, \quad e_j = -\frac{1}{\ln(n)} \sum_{i=1}^{n} p_{ij} \ln(p_{ij})$$ (4)

Step 3: Information utility and weight calculation. The information utility value is obtained through information entropy, and then the weight is obtained through normalization.

$$d_j = 1 - e_j, \quad w_j = \frac{d_j}{\sum_{j=1}^{m} d_j}$$ (5)

3.3 Objective Evaluation - CRITIC Weighting Method

The CRITIC weighting method is an objective weighting method based on data volatility, and its advantage lies in the comprehensive measurement of volatility and conflict, while taking into account the variability of indicators and the correlation of indicators. This method can use the objective properties of data for scientific evaluation.

Step 1: Indicator forwardization and data standardization. All indicators are converted into positive indicators and dimensionless.

Step 2: Indicator volatility and conflict. Volatility $S$ is expressed in terms of standard deviation, and larger standard deviations indicate greater volatility and higher weights. Conflicting $R$ is expressed using the correlation coefficient, and the larger the correlation coefficient value between indicators, the less conflicting and the lower the weight.

$$S_j = \sqrt{\frac{\sum_{j=1}^{n} (z_{ij} - \bar{z}_j)^2}{n-1}}, \quad R_j = \sum_{i=1}^{m} (1 - r_{ij})$$ (6)

Step 3: Indicator information content and weight. The information content $C$ is expressed by the product of volatility and conflict, and the weight is obtained by normalizing the amount of information.

$$C_j = S_j \times R_j, \quad w_j = \frac{c_j}{\sum_{j=1}^{m} c_j}$$ (7)

3.4 Objective Evaluation - Fuzzy Synthetic Evaluation

Fuzzy synthetic evaluation method is a comprehensive evaluation method based on fuzzy mathematics, which transforms qualitative evaluation into quantitative evaluation according to the affiliation theory of fuzzy mathematics. The model can make an overall evaluation of the objects affected by multiple factors. This method has the characteristics of clear results and strong systematic, which can better solve vague and difficult to quantify problems, and is suitable for solving various non-deterministic problems.
Step 1: Synthetic evaluation factor set and evaluation set. Factor set $U$ is a set of various indicators that affect the evaluation object. Evaluation set $V$ is a set of various possible outcomes for the evaluation object.

$$U = (u_1, u_2, \ldots, u_m), \ V = (v_1, v_2, \ldots, v_n)$$ (8)

Step 2: Fuzzy synthetic evaluation matrix and factor weight vector. Through the membership degree $r$ from the factor set to the evaluation set, the fuzzy synthetic evaluation matrix is obtained. The weight $A$ of each factor is used to form a set of weights.

$$R = (r_{ij})_{m \times n} = \begin{bmatrix} r_{11} & \cdots & r_{1n} \\ \vdots & \ddots & \vdots \\ r_{m1} & \cdots & r_{mn} \end{bmatrix}$$ (9)

$$A = (a_1, a_2, \ldots, a_m)$$ (10)

Step 3: Fuzzy vector and synthetic evaluation model. Change the fuzzy vector $A$ on $U$ to the fuzzy vector $B$ on $V$ by fuzzy change. The column vector $C$ with respect to the parameter specified for each class $v$. The system score $S$ is the matrix product of the fuzzy vector $B$ and the parameter column vector $C$.

$$B = A \ast R, C = (c_1, c_2, \ldots, c_n)^T, S = B \ast C$$ (11)

3.5 Subjective And Objective Fusion Evaluation Model

On the one hand, the evaluation model obtains the dimension weights through the AHP method based on expert scoring. On the other hand, based on the real data collected by enterprises, an integrated evaluation method including entropy weight method, CRITIC method and fuzzy synthetic evaluation method is designed to obtain the indicator weights. Finally, the subjective evaluation results and objective evaluation results are combined to obtain the comprehensive evaluation weight results.

Figure 4. Subjective and objective synthetic evaluation process
4 EXPERIMENTAL RESULTS

4.1 Subjective Evaluation Experimental Results

Based on expert scoring and AHP method, the importance of the three dimensions of AOE model is compared and the comprehensive weight calculation is carried out. The result of the experiment is that the capability weight is 30%, the operation weight is 50%, and the effectiveness weight is 20%.

**TABLE I. AHP JUDGMENT MATRIX**

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Ability</th>
<th>Operation</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability</td>
<td>1</td>
<td>0.5</td>
<td>2</td>
</tr>
<tr>
<td>Operation</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Effect</td>
<td>0.5</td>
<td>0.5</td>
<td>1</td>
</tr>
</tbody>
</table>

**TABLE II. AHP ANALYSIS RESULTS**

<table>
<thead>
<tr>
<th>Item</th>
<th>Eigenvector</th>
<th>Weight</th>
<th>Maximum Eigenvalue</th>
<th>CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability</td>
<td>0.936</td>
<td>31.19%</td>
<td>3.054</td>
<td>0.027</td>
</tr>
<tr>
<td>Operation</td>
<td>1.471</td>
<td>49.05%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effect</td>
<td>0.593</td>
<td>19.76%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**TABLE III. CONSISTENCY TEST RESULTS**

<table>
<thead>
<tr>
<th>Maximum Root</th>
<th>CI</th>
<th>RI</th>
<th>CR</th>
<th>Consistency Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.054</td>
<td>0.027</td>
<td>0.52</td>
<td>0.052</td>
<td>pass</td>
</tr>
</tbody>
</table>

**Figure 5.** Three evaluation dimensions weight results based on AHP

The expert scoring method based on the weighted evaluation method is used to score the secondary indicators in the AOE model, and the results are shown in the subscript.
### TABLE IV. EXPERT SCORING EVALUATION RESULTS

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Key indicators</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability evaluation</td>
<td>Product</td>
<td>10%</td>
</tr>
<tr>
<td></td>
<td>Organization</td>
<td>15%</td>
</tr>
<tr>
<td></td>
<td>Channel</td>
<td>15%</td>
</tr>
<tr>
<td></td>
<td>Platform</td>
<td>20%</td>
</tr>
<tr>
<td></td>
<td>Resource</td>
<td>20%</td>
</tr>
<tr>
<td></td>
<td>Assessment</td>
<td>20%</td>
</tr>
<tr>
<td>Operation evaluation</td>
<td>Digital Village</td>
<td>20%</td>
</tr>
<tr>
<td></td>
<td>Smart Community</td>
<td>20%</td>
</tr>
<tr>
<td></td>
<td>Street Shops</td>
<td>15%</td>
</tr>
<tr>
<td></td>
<td>Corporate Membership</td>
<td>15%</td>
</tr>
<tr>
<td></td>
<td>Specific Business</td>
<td>15%</td>
</tr>
<tr>
<td></td>
<td>Commercial Buildings</td>
<td>10%</td>
</tr>
<tr>
<td></td>
<td>Campus Market</td>
<td>10%</td>
</tr>
<tr>
<td>Effect evaluation</td>
<td>Multi-business Customers</td>
<td>40%</td>
</tr>
<tr>
<td></td>
<td>Customer Value</td>
<td>30%</td>
</tr>
<tr>
<td></td>
<td>Customer Loyalty</td>
<td>30%</td>
</tr>
</tbody>
</table>

#### 4.2 Entropy Weight Method Experimental Results

Based on the entropy weight method, the weights of each indicators are calculated, and the weight coefficients of each indicators in the three dimensions are shown in the following table.

### TABLE V. ABILITY: ENTROPY METHOD RESULTS

<table>
<thead>
<tr>
<th>Item</th>
<th>Information entropy value</th>
<th>Information utility value</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Product</td>
<td>1</td>
<td>0</td>
<td>0.00%</td>
</tr>
<tr>
<td>Organization</td>
<td>0.9933</td>
<td>0.0067</td>
<td>2.14%</td>
</tr>
<tr>
<td>Channel</td>
<td>0.9599</td>
<td>0.0401</td>
<td>12.80%</td>
</tr>
<tr>
<td>Platform</td>
<td>0.87</td>
<td>0.13</td>
<td>41.46%</td>
</tr>
<tr>
<td>Resource</td>
<td>0.9933</td>
<td>0.0067</td>
<td>2.14%</td>
</tr>
<tr>
<td>Assessment</td>
<td>0.87</td>
<td>0.13</td>
<td>41.46%</td>
</tr>
</tbody>
</table>

### TABLE VI. OPERATION: ENTROPY METHOD RESULTS

<table>
<thead>
<tr>
<th>Item</th>
<th>Information entropy value</th>
<th>Information utility value</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digital Village</td>
<td>0.9376</td>
<td>0.0624</td>
<td>9.92%</td>
</tr>
<tr>
<td>Smart Community</td>
<td>0.8691</td>
<td>0.1309</td>
<td>20.82%</td>
</tr>
</tbody>
</table>
### TABLE VII. EFFECT: ENTROPY METHOD RESULTS

<table>
<thead>
<tr>
<th>Item</th>
<th>Information entropy value</th>
<th>Information utility value</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multi-business Customers</td>
<td>0.993</td>
<td>0.007</td>
<td>31.70%</td>
</tr>
<tr>
<td>Customer Value</td>
<td>0.9876</td>
<td>0.0124</td>
<td>56.61%</td>
</tr>
<tr>
<td>Customer Loyalty</td>
<td>0.9974</td>
<td>0.0026</td>
<td>11.69%</td>
</tr>
</tbody>
</table>

### 4.3 CRITIC Method Experimental Results

Based on the CRITIC method, the weight of each indicator is calculated, and the weight coefficients of each indicator in the three dimensions are obtained as shown in the following table.

### TABLE VIII. ABILITY: CRITIC METHOD RESULTS

<table>
<thead>
<tr>
<th>Item</th>
<th>Indicator variability</th>
<th>Indicator conflict</th>
<th>Information content</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Product</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0.00%</td>
</tr>
<tr>
<td>Organization</td>
<td>0.148</td>
<td>2.372</td>
<td>0.35</td>
<td>8.93%</td>
</tr>
<tr>
<td>Channel</td>
<td>0.299</td>
<td>2.122</td>
<td>0.635</td>
<td>16.19%</td>
</tr>
<tr>
<td>Platform</td>
<td>0.447</td>
<td>2.372</td>
<td>1.061</td>
<td>27.07%</td>
</tr>
<tr>
<td>Resource</td>
<td>0.148</td>
<td>5.507</td>
<td>0.813</td>
<td>20.74%</td>
</tr>
<tr>
<td>Assessment</td>
<td>0.447</td>
<td>2.372</td>
<td>1.061</td>
<td>27.07%</td>
</tr>
</tbody>
</table>
TABLE IX. OPERATION: CRITIC METHOD RESULTS

<table>
<thead>
<tr>
<th>Item</th>
<th>Indicator variability</th>
<th>Indicator conflict</th>
<th>Information content</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digital Village</td>
<td>0.821</td>
<td>3.992</td>
<td>3.277</td>
<td>12.13%</td>
</tr>
<tr>
<td>Smart Community</td>
<td>0.905</td>
<td>4.828</td>
<td>4.371</td>
<td>16.18%</td>
</tr>
<tr>
<td>Street Shops</td>
<td>0.692</td>
<td>7.097</td>
<td>4.911</td>
<td>18.18%</td>
</tr>
<tr>
<td>Corporate Membership</td>
<td>2.155</td>
<td>5.029</td>
<td>10.839</td>
<td>40.13%</td>
</tr>
<tr>
<td>Specific Business</td>
<td>0.27</td>
<td>4.529</td>
<td>1.223</td>
<td>4.53%</td>
</tr>
<tr>
<td>Commercial Buildings</td>
<td>0.124</td>
<td>5.398</td>
<td>0.67</td>
<td>2.48%</td>
</tr>
<tr>
<td>Campus Market</td>
<td>0.234</td>
<td>7.357</td>
<td>1.719</td>
<td>6.36%</td>
</tr>
</tbody>
</table>

TABLE X. EFFECT: CRITIC METHOD RESULTS

<table>
<thead>
<tr>
<th>Item</th>
<th>Indicator variability</th>
<th>Indicator conflict</th>
<th>Information content</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multi-business Customers</td>
<td>0.064</td>
<td>1.261</td>
<td>0.081</td>
<td>16.55%</td>
</tr>
<tr>
<td>Customer Value</td>
<td>0.185</td>
<td>1.899</td>
<td>0.35</td>
<td>71.39%</td>
</tr>
<tr>
<td>Customer Loyalty</td>
<td>0.061</td>
<td>0.971</td>
<td>0.059</td>
<td>12.06%</td>
</tr>
</tbody>
</table>

4.4 Experimental Results of Fuzzy Synthetic Evaluation Method

Based on the fuzzy synthetic evaluation method, the weight of each indicator is calculated, and the weight coefficient of each indicator in the three dimensions is obtained as shown in the following table.

TABLE XI. ABILITY: FUZZY SYNTHETIC EVALUATION RESULTS

<table>
<thead>
<tr>
<th>Item</th>
<th>Degree of affiliation</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Product</td>
<td>0.204</td>
<td>20%</td>
</tr>
<tr>
<td>Organization</td>
<td>0.182</td>
<td>18%</td>
</tr>
<tr>
<td>Channel</td>
<td>0.147</td>
<td>15%</td>
</tr>
<tr>
<td>Platform</td>
<td>0.137</td>
<td>14%</td>
</tr>
<tr>
<td>Resource</td>
<td>0.192</td>
<td>19%</td>
</tr>
<tr>
<td>Assessment</td>
<td>0.137</td>
<td>14%</td>
</tr>
</tbody>
</table>

TABLE XII. OPERATION: FUZZY SYNTHETIC EVALUATION RESULTS

<table>
<thead>
<tr>
<th>Item</th>
<th>Degree of affiliation</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digital Village</td>
<td>0.164</td>
<td>16%</td>
</tr>
<tr>
<td>Smart Community</td>
<td>0.119</td>
<td>12%</td>
</tr>
</tbody>
</table>
### TABLE XIII. EFFECT: FUZZY SYNTHETIC EVALUATION RESULTS

<table>
<thead>
<tr>
<th>Item</th>
<th>Degree of affiliation</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multi-business Customers</td>
<td>0.185</td>
<td>19%</td>
</tr>
<tr>
<td>Customer Value</td>
<td>0.484</td>
<td>48%</td>
</tr>
<tr>
<td>Customer Loyalty</td>
<td>0.331</td>
<td>33%</td>
</tr>
</tbody>
</table>

### 4.5 Experimental Results of Subjective And Objective Fusion Evaluation Method

The weight calculation results of the subjective and objective fusion evaluation method and the scoring rules of each indicator are shown in the following table.

### TABLE XIV. INTEGRATED EVALUATION RESULTS

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Key indicators</th>
<th>Weight</th>
<th>Scoring rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability evaluation</td>
<td>Product</td>
<td>9%</td>
<td>Score by completion ratio:[0,30,60,100]</td>
</tr>
<tr>
<td></td>
<td>Organization</td>
<td>13%</td>
<td>Score by completion ratio:[0,30,60,100]</td>
</tr>
<tr>
<td></td>
<td>Channel</td>
<td>15%</td>
<td>Score by completion ratio:[0,30,60,100]</td>
</tr>
<tr>
<td></td>
<td>Platform</td>
<td>22%</td>
<td>Score by completion ratio:[0,30,60,100]</td>
</tr>
<tr>
<td></td>
<td>Resource</td>
<td>18%</td>
<td>Score by completion ratio:[0,30,60,100]</td>
</tr>
<tr>
<td></td>
<td>Assessment</td>
<td>22%</td>
<td>Score by completion ratio:[0,50,100]</td>
</tr>
<tr>
<td>Operation</td>
<td>Digital Village</td>
<td>18%</td>
<td>Scoring is linear within 100%</td>
</tr>
<tr>
<td>evaluation</td>
<td>Smart Community</td>
<td>19%</td>
<td>Scoring is linear within 100%</td>
</tr>
<tr>
<td></td>
<td>Street Shops</td>
<td>4%</td>
<td>Scoring is linear within 100%</td>
</tr>
</tbody>
</table>
### 5 MODEL APPLICATION AND EMPIRICAL ANALYSIS

Based on this model, a data table tracking mechanism is established to regularly collect key capacity building progress, integrated operation development data, and fusion effect data in the process of integrated development. It is used to understand the specific situation of each province and each link of integrated development. Through the data table, comprehensive evaluation can be realized, including the scoring of various indicators of provincial companies, the scoring of three dimensions, the comprehensive score of integration, and the ranking. Based on the summary of provinces, the overall progress of integration capacity building, the total progress of integration operation, and the total effectiveness of integration can be comprehensively analyzed.

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Weighting</th>
<th>Scoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corporate Membership</td>
<td>20%</td>
<td>Linear within 100%</td>
</tr>
<tr>
<td>Specific Business</td>
<td>12%</td>
<td>Linear within 100%</td>
</tr>
<tr>
<td>Commercial Buildings</td>
<td>9%</td>
<td>Linear within 100%</td>
</tr>
<tr>
<td>Campus Market</td>
<td>10%</td>
<td>Linear within 100%</td>
</tr>
<tr>
<td>Effect evaluation</td>
<td>20%</td>
<td>Linear within 100%</td>
</tr>
<tr>
<td>Multi-business Customers</td>
<td>35%</td>
<td>Linear within 100%</td>
</tr>
<tr>
<td>Customer Value</td>
<td>39%</td>
<td>Linear within 100%</td>
</tr>
<tr>
<td>Customer Loyalty</td>
<td>27%</td>
<td>Linear within 100%</td>
</tr>
</tbody>
</table>

**Figure 6.** Evaluation model application on some provincial corporations
Figure 7. Evaluation results for overall integration development level

The synthetic evaluation conclusion of integration development is as follows. First, integration capacity building is the foundation, and the construction progress is considerable. The development of products is relatively mature, some provincial companies need to be improved in terms of organization and resources, and some provincial companies need to strengthen construction in terms of contacts, platforms and assessments. Second, integration operation has achieved phased results, and some scenarios need to be continuously expanded. Scenarios such as digital villages and smart communities have achieved good operational results driven by the resources of the conference, while scenarios such as campus markets and pendant markets still need to be continuously expanded. Third, the integration effect is significant and the value and loyalty of integrated customers both have been improved. The penetration rate of integrated customers is 30%+, and there is still room for improvement in the scale of integrated customers.
Empirical analysis results show that customer integration development can promote customer value and loyalty. First, the integration of H and C areas can increase value by 77% and loyalty rate by 48%. Compared with C+B integration, C+H integration is more significant in value enhancement. Second, the integration of B and C can increase value by 29% and loyalty rate by 52%. Compared with C+H integration, C+B integration is more significant in terms of loyalty improvement. Third, C+H+B full integration has the highest value improvement rate of 92% and the highest loyalty promotion rate of 68%, and the development towards full integration has achieved remarkable results.

6 CONCLUSIONS

This paper constructs a subjective and objective comprehensive evaluation model for the level of enterprise integration development, which can fully integrate the management experience of
experts and the information behind the statistical data. Through empirical analysis, it is proved that the model can effectively reflect the overall progress of enterprise integration development, and at the same time reflect more mature aspects and relatively backward aspects, thus assisting enterprises to analyze problems in the process of integration development. Furthermore, the model achieves regional evaluation by scoring, thus achieving the comparison of the integration development level of enterprises in different regions, and successfully locates the problems in various regions through radar map analysis. To sum up, the evaluation model can effectively assist the scientific decision-making in the management and operation process of enterprise integration development.

REFERENCES

[10] Pang, 2016, Jiuyu shangquan TOPSIS fa de sanwang ronghe fazhan pinggu moxing ji shizheng yanjiu [Development evaluation model and empirical study of triple play based on entropy-weight TOPSIS method], Beijing University of Post and Telecommunications, 2016.12
Decision Tree-Based Human Resources Forecasting and Enterprise Project Management
COVID-19 Impact and Enterprise Response

Xue Zhao1, 2, a, Bin Zhao3, b, Zhenjun Du3, 4, Fengshan Zou3, 4

technzhaoxue@gmail.coma, zhaobin@stumail.neu.edu.cnb

Daniel L. Goodwin College of Business, Benedictine University, Chicago, 60601, the United States1
Shenyang packaging materials Co. Ltd, Shenyang, 110032, China2
SIASUN Robot & Automation Co., Ltd., Shenyang, 110168, China3
Shenyang Institute of Automation, Chinese Academy of Science, Shenyang, 110016, China4

Abstract: Under the impact of the COVID-19 epidemic, it has had an enormous impact on Chinese high-tech enterprises. Based on the practical application of SIASUN's project management and human resources, this paper discusses a series of solutions by studying the current situation of SIASUN's project management system and the problems existing in the actual needs of employees. Scientific project management and human resource concept will promote the project's development. This paper expounds on project management and human resources universality and timeliness. To achieve technological innovation of project management and human resources, production capacity innovation, and management system innovation to improve the core competitiveness of enterprises in the large national market environment. Finally, under the impact of COVID-19, this paper expounds on some problems of project management and human resources in high-tech enterprises. This paper analyses the problems of project management and human resources and gives a systematic solution. The decision tree model is used to predict employee turnover and ensure the smooth progress of the project. Enterprises need to complete more projects in a limited time, ensure the success rate of projects and the effective use of resources, reduce production costs, and significantly improve sales and market share.

Keywords: Project Management, Decision Tree, Human Resources, High-Tech Enterprises.

1 INTRODUCTION

With the development of the economy and the change in environment, the increasingly fierce global competition, the increasing diversity and uncertainty of product types, and user needs, the competition between enterprises has gradually turned into the competition between project management and human resources [4, 8]. Under the impact of COVID-19, the traditional project management and human resources problems increasingly exposed their inherent two defects: first, the project management and human resources problems needed to be more comprehensive, simple, and less advanced technologies and methods were used [1]. Not only did the cost of project management and human resources remain high, but also the slow progress of the project led to project failure. Second, project management and human resource management are
independent, which increases the overall management cost of the project and human resource management and reduces its overall competitive advantage. Therefore, management and human resource issues are essential to meet enterprises' needs, safeguard enterprises' interests, and constantly improve project management's safety for management control in all aspects of the project.

Under the impact of the COVID-19 epidemic, the international situation is grim [9]. Although various countries are opening up one after another, the epidemic's effect continues [2, 6, 10]. Some enterprises, in some aspects of the operation of the thought, only stay in the project, and human resource management, on the surface of understanding, can only be from a few elements, profound multi-level management knowledge. How to combine human resource management and project management organically has become the main problem to be solved in project design and human resource management. From the perspective of an enterprise, the employee turnover rate is an issue that must be paid attention to by the enterprise. Various factors affect employee turnover (salary, travel, work environment satisfaction, work engagement, overtime, promotion, salary increase ratio, etc.) and the corresponding records of whether an employee has resigned. This paper will establish a decision tree model to predict the employee's likelihood to leave from the factors that affect employee resignation and whether or not the employee resigns to increase the project success rate. Each sample has a set of attributes and a classification result. That is, the classification result is known. The decision tree can obtain by learning these samples, giving the correct classification and regression for the new data.

2 PROJECT MANAGEMENT THEORY

As a unique feature of R&D projects, project management has a crucial impact on human resource management [3]. In a project-oriented organization, management support plays a leading role [7]. In project management, the project usually needs to improve at taking care of employees because of the project's progress. The demand for profits and the response to customer needs often precede employee benefits. Employees in the project often direct feedback on the project's negative impact because their wages and benefits cannot be satisfied. The pressure of the project directly affects employees, and employees' expectations cannot be met, which also affects the project. Human resource management uses salary dispersion, training, career promotion, and communication to regulate the interaction of organizational management.

2.1 Universality and Clarity

The project and human resource management links include purchase, production, material management, marketing, logistics customer demand, capital flow, information flow, inventory cost, employee recruitment, and employee training [5]. Figure 1 shows the main management elements of the project management life cycle. For a long time, the study of project management was mainly carried out at the general academic research level rather than guiding the practice of enterprises. Project and human resource management in high-tech enterprises have universality and clarity. The concept of project management is: comprehensively and effectively manage information flow, capital flow, and logistics, including planning, coordination, organization, and control.
The control of the project management system involved in the project and human resource management theory aims to enable enterprises to achieve customer products and optimize the control of existing processes. Among them, optimizing the total cost is particularly important because the cost is the essential element for an enterprise to survive. The project and human resource management theories not only emphasize the concept and methods of the comprehensive management model but also organically combine all aspects of the project and human resource management, which can achieve the highest efficiency of project management. Enterprises must find the correct positioning. If an enterprise does the project behind closed doors, it is difficult to cover all aspects of the project, eventually leading to project delay or even project failure. Because of this, project and human resource management teach managers to integrate resources and monitor projects using rational and scientific management methods.

2.2 Timeliness and Rigor

High-tech enterprises use advanced project and human resource management theory to organize their enterprises. In the whole enterprise management process, timeliness and rigor are very important for the decision-makers and managers of enterprises. The traditional management mode has many shortcomings, such as poor confidentiality, low efficiency, and high cost. A long time will produce a large number of documents and data, which has brought many difficulties to the project's implementation and maintenance. Because of this, high-tech enterprises must establish a complete project and human resources management system. This paper focuses on the design and implementation of Siasun's project and human resource management.

Because of the timeliness and rigor of project management, high-tech enterprises in many industries and regions have been widely recognized and developed. For any project task in the process of execution, there are specific start and end times. Therefore, the implementation goal and time have been made clear before the project implementation. Resource sharing in the
information age has been popularized and widely implemented in high-tech enterprises. In order to ensure the accuracy and rigor of project and human resource management information, the effective implementation of information resource sharing in each link to achieve timely, accurate, and comprehensive protection of project and human resource management information.

3 DECISION TREE BASED HUMAN RESOURCES FORECASTING

The factors influencing the prediction of human resource demand are as follows: (1) the enterprise's production and operation tasks and its demand for human resources in a certain period in the future. (2) The expected employee turnover rate and job vacancy size. (3) The influence of the improvement of production technology and the change of organizational management mode on demand for human resources. (4) The influence of an enterprise's decision to improve product or service quality or enter a new market on human resource demand. (5) Constraints of enterprise's financial resources on human resource demand. This paper studies the decision tree training and testing of expected employee turnover.

3.1 An Overview of Human Resource Demand Forecasting

In order to ensure the smooth progress of the project, it is necessary to forecast employee turnover. The corresponding record of whether an employee has resigned is used to improve the data. The data set is divided into training sets and test sets. Training stage: Construct a tree from a training set (special features from the heel node and how to carry out feature segmentation). Test stage: Walk from top to bottom according to the constructed tree model. Once the decision tree is constructed, the task of classification or prediction is effortless and only needs to walk once. Then the difficulty lies in how to construct a tree. Therefore, the generation of a decision tree mainly consists of the following two steps: 1. Node splitting: Generally, when the attributes represented by a node cannot be judged, the node will be divided into two child nodes (if it is not a binary tree, it will be divided into n child nodes). The permutation candidate threshold conditions are compared, and the candidate condition with the lowest entropy is selected as the parent node for tree generation; 2. Threshold determination: Select an appropriate threshold to minimize the classification error rate. In order to ensure the smooth progress of the project, it is necessary to forecast employee turnover. The corresponding record of whether an employee has resigned is used to improve the data. The data set is divided into training sets and test sets. If the decision tree depends on the mathematical calculation method, it can achieve more ideal results. The mathematical expression is as follows:

\[ (x,y) = (x_1, x_2, x_3, ..., x_k, y) \]  

The relevant variable \( y \) indicates that we are trying to understand, classify, or more general results. Other variables, \( x_1, x_2, x_3, ..., x_k \), etc., help us achieve our goals.

3.2 Decision Tree Training and Testing

1. Information entropy:
The classification standard used by ID3 is information gain (Entropy), which represents the degree to which the uncertainty of the sample set is reduced when the information of feature $A$ is known. Information entropy of data set:

$$Entropy = H(D) = -\sum[p(x_i) \cdot \log_2(P(x_i))] = -\sum_{k=1}^{K} \frac{|C_k|}{|D|} \log_2 \frac{|C_k|}{|D|}$$  \hspace{1cm} (2)

Where $C_k$ refers to the subset of samples belonging to class $k$ in set $D$. Therefore, when the information entropy is at its maximum of 1, the classification effect is the worst; when the information entropy is at its minimum of 0, it is the state of complete classification. Because entropy equals zero is the ideal state, and in practice, entropy is between zero and one. For a feature $A$, the conditional entropy $H(D \mid A)$ of dataset $D$ is:

$$H(D \mid A) = \sum_{i=1}^{n} \frac{|D_i|}{|D|} H(D_i)$$

$$= -\sum_{i=1}^{n} \frac{|D_i|}{|D|} (\sum_{k=1}^{K} \frac{|D_{ik}|}{|D_i|} \log_2 \frac{|D_{ik}|}{|D_i|})$$  \hspace{1cm} (3)

Where $D_i$ represents the sample subset of the i-th value of feature $A$ in $D$, $D_{ik}$ stands for $D_i$ and is a subset of samples belonging to class $k$.

2. Information gain:

Information gain = information entropy - conditional entropy:

$$Gain(D, A) = H(D) - H(D \mid A)$$  \hspace{1cm} (4)

The greater the information gain is, the more significant the purity improvement obtained by using feature $A$ to divide. The classification of training data can reach 0 error rate, but because the new data is different from the training data, the error rate increases in the face of the new data. The decision tree is to obtain the statistical information of the data by analyzing the training data rather than being tailored to the training data.

\[\text{Figure 2: Decision tree prediction.}\]
Therefore, in order to avoid too fine segmentation, C4.5 improves ID3. In C4.5, the optimization term should be divided by the cost of too-fine segmentation, called the information gain rate. If the segmentation is too fine, the information gain rate will decrease. Other than that, the principles are the same as ID3. Figure 2 shows that the turnover rate of frequent business trips is the highest, and the number of research departments accounts for the most significant proportion. However, the turnover of personnel in the sales department is higher, and the turnover rate is higher. The turnover rate of men and single employees is relatively high. Respect for employees is the most fundamental principle if enterprises want to retain talent.

4 PROJECT AND HUMAN RESOURCE MANAGEMENT ISSUES

4.1 Problem of Technical Requirements

For the enterprise, there will be mistakes if there are many projects, and everyone will focus on the project differently. Only by discovering the actual needs behind the project can we take timely actions before problems occur and significantly increase the efficiency of successful development. For high-tech enterprises, there are very few experienced and skilled people, so it is impossible to have such people in every link. The project management of high-tech enterprises often involves many complex and diverse technologies. In the past, it took much work for technical developers to refine the project's technical unit, summarize the technical content required by each part, draw a flow chart and arrange it for the designated personnel. Generally speaking, technical personnel will analyze where they want to go, and the project will be stalled when they encounter technical problems.

4.2 Problem of Team Structure

Under the impact of COVID-19, it is necessary to conduct a comprehensive analysis of enterprise projects and human resource management to realize the construction of efficient production system management. Typically, inexperienced project managers go by the seat of their pants when given a project assignment. Team structure refers to the composition of team members, which is the basis of team cooperation and plays an essential role in the success of projects. The project team members may come from different departments, so they must return to the original department after the project. This means the team needs more tacit understanding and needs to be reconciled as soon as possible. However, as the project progresses, new members may be added to the project, and the project's structure should be adjusted slightly, which often makes it difficult for the project executive to integrate the team structure.

4.3 Problem of Tasks Division

Under the impact of COVID-19, project management and human resources of enterprises are more prominent. Scientific project management concepts will promote the project's development, but if there are mechanical implementation and severe compliance problems in the process of project management. Employees do not express their expectations during project development because many people need to understand their abilities. Therefore, it is difficult in project development to assign tasks according to workability, give play to personal strengths, and implement responsibilities to everyone. Task allocation is more challenging to balance and
care about the ideas of project members, etc. In the whole process, project and human resource management generally use scientific management mode to make project executives more aligned with the process. Not only can it provide a complete product research and development process, but it also can provide a basis for purchasing, production, and sales. The goal of project management is success, and successful goals need to meet three criteria: business goals, cost, and schedule.

5 PROBLEM AVOIDANCE METHODS

5.1 Strengthen Risk and Quality Management

Due to the complexity of the project, the project risk resulted in many uncertain factors. The problem of project risk occurrence often has hidden risks and uncertainties because it is difficult to predict risks in all types of project links. However, preventive measures are often taken in a typical project risk control, which requires an enterprise's technology accumulation. Only good technology accumulation and problem-solving mechanisms can be more effective in prevention. Therefore, to achieve the work goal, it is necessary to divide the specific tasks so that the tasks can be carried out in parallel. By clarifying the tasks, we can determine the project team and make a good division of labor. Project risk management is a comprehensive risk management project. The goals and tasks of projects are often fuzzy boundary conditions in the development of robotics projects, often to change the specific Customer needs to adjust the program.

Based on ensuring the interests of enterprises, the higher the quality, the stronger the competitive advantage, the better the reputation, and ultimately improve the visibility and influence of enterprises. Quality management research is a solid guarantee for production projects to enter the market. Under the requirement of quality management, delivering deliverables and evaluating the quality of activities are not conducive to improving the quality of engineering projects but also to implementing remedial work plan deficiencies. Under the premise of closed-loop control and process control, the overall quality can be guaranteed. The implementation of monitoring management can significantly improve the quality and enhance the competitive advantage of enterprises in the market.

5.2 Incentive Management of Human Resource

For project members to carry out the project with passion and not dull, a fundamental premise is to trust the project members fully. In real life, the project manager often likes to ask questions and ask the project members to explain the project's progress to him, which will cause the employees to be deeply distrusted. Project members need to be recognized, and they should also be given some practical rewards. The results in Figure 3 show that the higher the satisfaction with the job reward provided, the higher the job satisfaction. In addition, when project managers were more satisfied with their company's work rewards, they also had higher satisfaction with their company's HRM practices. The more satisfied managers were with their job rewards, promotion opportunities, benefits packages, and total compensation packages, the higher their job satisfaction. This indicates a significant correlation between project managers' satisfaction with their company's HRM practices and their overall job satisfaction.
Enterprises put more energy and money into the talent training and management of high-tech enterprises. The paper puts forward higher requirements for high-tech enterprises to achieve their ultimate goals. In project management, the project manager should not only motivate the project members positively but also take some harmful incentive methods to exert a little work pressure on the project members. Many successful managers are good at using this psychology of their subordinates. The potential of employees is also realized in this way.

5.3 Project control and Closure

The famous milestone plan is the classic clever method of project control. This is a goal, which is a set of activities to achieve the project goals. Assign tasks, organize team, and determine project objectives and scope. Make and review project plans, track, supervise and control the project according to the Project Plan, and ensure the smooth implementation and completion of the project according to the plan. Milestone planning aims to control the project's progress and ensure the achievement of the overall goal by establishing milestones and verifying the achievement of each milestone.

At the end of the project, the project manager should identify the success indicators of the project from multiple perspectives and self-determine whether the project is successful or not. Generally speaking, project management will consider the problem from multiple perspectives in the end stage: finance, time, quality, human resources, environment, project planning, project control, and other aspects to measure the completion of the project. Although different people have their standards for the results of the project, in the final analysis, we can be summarized as follows: (1) to achieve the project expectations; (2) to evaluate the effectiveness of the project; (3) to customer's evaluation of the use effect.
6 CONCLUSION AND PROSPECT

Under the impact of COVID-19, project and human resource management are to effectively manage the whole project process and human resource management using systematic theories and methods. Because the projects of high-tech enterprises are characterized by uncertainty, uniqueness, and urgency, it is necessary to have advanced management software and management means when managing projects and human resources. Based on the project and human resource management problems involved in the technology development process of SIASUN Company, this paper starts with the current situation, existing problems, and actual needs of management. This paper analyzes the current situation and some problems of high-tech enterprises’ projects and human resource management. It can make the project more efficient based on strengthening risk management, quality management, human resource incentive, project management control, building an advanced platform, human resource early warning, strategic resource allocation, and scientific management technology. The decision tree model is used to predict employee turnover and ensure the smooth progress of the project. It is challenging to find a perfect solution to a real problem. It is of great significance to improve the speed and reasonable development of the project and human resource management for the operation efficiency of enterprises.

REFERENCES

A Study of Remanufacturing Pricing Decisions Considering Recycling Quality and Retailer CSR Inputs

Wensheng Yang¹,a, Anxin Hu¹,b*
yangwsh@126.com,a, *839354906@qq.comb

School of Economics and Management (School of Tourism), Dalian University, No. 10, Xuefu Street, Jinzhou New District, Dalian 116622, China¹

Abstract: In order to improve the sales and production of recycled products, a closed-loop supply chain model of manufacturer recycling is constructed based on the differences in retailers' CSR inputs and the quality level of recycled products, and the optimal pricing of the remanufacturing supply chain is investigated through a Stackelberg game. The results show that supply chain revenue increases when consumers' willingness to pay for remanufactured products approaches that of new products. From the perspective of expanding the market demand for remanufactured products and increasing the supply chain revenue, the higher quality of returned products and consumer CSR sensitivity are better, while the quality of recyclable products also affects the retailer CSR input level, and the efficiency of CSR input and the quality of recycled waste products can be considered to enhance the economic and environmental benefits.

Keywords: Recycling Quality, Consumer Preference, Corporate Social Responsibility, Stackelberg Game.

1 INTRODUCTION

In recent years, due to a series of problems such as epidemic, shortage of production raw materials and environmental pollution, enterprises pay more attention to the recycling and reuse of used products, and scholars gradually regard the study of remanufacturing closed-loop supply chain as one of the research hotspots. In reality, due to the different usage habits and degrees of use in the recycling process, there is uncertainty about the participation value of recycled products, and consumers are more inclined to purchase new products than remanufactured products. Therefore, in order to increase consumers' attention to recycled products and alleviate the shortage of production resources, it is important to study the closed-loop supply chain with the difference in CSR input and recycling quality of retailers under the consumer preference.

Based on the above problems, the review of relevant literature shows that even if the quality of new products and remanufactured products is the same, consumers will have different pricing for new products and remanufactured products due to different consumer cognition. Zhao et al. introduced consumer environmental preferences to study the pricing decisions of supply chain members and the joint decision problem of subsidy shares between remanufacturers and consumers [11]. Hong et al. concluded from their study that the degree of consumer preference...
for remanufactured products affects the price of remanufactured products at wholesale, retail and recycling, and that the recycling price increases with the increase in wholesale and retail prices [3]. Sun et al. develop but-channel and dual-channel closed-loop supply chain models based on consumer preferences and consumer equity concerns to discuss the impact of consumer acceptance of remanufactured goods and equity concerns on members’ decisions and profits [9]. For the quality of raw materials in the recycling process, Feng et al. construct a multi-level competitive recycling and remanufacturing supply chain game model with two manufacturers and multiple recyclers to analyze the pricing coordination problem [2]. Taleizadeh et al. analyzed the dynamic pricing and recycling strategies of remanufacturers using a modal interval algorithm considering the uncertainty of scrap product quality [10].

Some scholars believe that donating a certain amount of money for public welfare is also a way to make CSR investment, Muller et al. showed that when consumers are philanthropically socially conscious, an increase in the size of donations can increase consumer demand [7], Dey et al. considered the sustainability of remanufacturing and competition in the retail market, added socially responsible influences to the market demand, and found that remanufacturing is not only beneficial to the overall supply chain in terms of economic benefits, but also in terms of environmental sustainability [1].

In summary, this study differs from previous work in several ways. Firstly, in order to increase the sales volume of remanufactured products, the influences of consumer preference and CSR input for remanufactured products, namely social donation, on the pricing of remanufactured products and supply chain profits were considered. Secondly, the study found that there was a certain relationship between the quality of recycled products and CSR input. Finally, combining the three aspects of consumer preference, CSR input of enterprises and the quality of recycled products, it provided a certain theoretical reference for alleviating the plight of raw materials and increasing the sales volume of remanufactured products.

2 SYMBOL DEFINITION AND HYPOTHESES

This section examines a closed-loop supply chain system for remanufacturing, taking into account differential pricing of retailers’ socially responsible inputs based on quality differences in recycled goods, as shown in Figure 1.

Based on the actual situation and relevant literature, the following assumptions are made:

Hypothesis 1: In the closed-loop supply chain, each member has completely symmetrical information and neutral preference for market risk. The Stackelberg game is dominated by the manufacturer, and the decision-making goal is to maximize their respective profits [5-6].
Hypothesis 2: Assume that only a single product category is considered, and the market is relatively mature, with enough old products available for recycling and remanufacturing, and consumers buy only one new product or remanufactured product\(^4\).

Hypothesis 3: Assumed that the new product and the remanufactured product have the same specifications, quality and function, but the manufacturing materials are different, so the consumer preferences for the new product and the remanufactured product are different\(^6\).

Hypothesis 4: Assumed that the quality of the recycled products can be remanufactured, the manufacturer passes the quality assessment, and the consumers have different recycling subsidies.

The meanings of relevant parameters of the model are shown in Table 1.

<table>
<thead>
<tr>
<th>Symbols</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_n, p_r )</td>
<td>The unit retail price of new/remanufactured products</td>
</tr>
<tr>
<td>( C_n )</td>
<td>Unit production cost of new products</td>
</tr>
<tr>
<td>( C_r )</td>
<td>Unit production cost of remanufactured products, ( C_r = C_n - C_d q ), ( C_d ) is the remanufacturing cost factor, expressed as the degree of influence of the quality of recycled goods on the remanufacturing cost</td>
</tr>
<tr>
<td>( w_n, w_r )</td>
<td>The unit wholesale price of new/remanufactured products</td>
</tr>
<tr>
<td>( e_r )</td>
<td>Retailer CSR investment level</td>
</tr>
<tr>
<td>( q )</td>
<td>Quality level of recycled waste product ( q \in (0,1) )</td>
</tr>
<tr>
<td>( B )</td>
<td>Recycling price for manufacturers to recycle waste products ( B = \sigma q, \sigma &gt; 0 ) indicates the maximum recovery price per unit that the manufacturer is willing to pay</td>
</tr>
<tr>
<td>( G )</td>
<td>Number of discarded products recycled by the manufacturer, ( G = h + kB ), ( h ) indicates the number of waste products that consumers participate in recycling without compensation, ( k &gt; 0 ) indicates the recovery price sensitivity factor</td>
</tr>
<tr>
<td>( \Pi^i_1 )</td>
<td>Profit of closed-loop supply chain member ( i ) in model ( j ), ( i = M, R; j = )</td>
</tr>
</tbody>
</table>

Assuming that the total market size is \( Q \), the willingness to pay for new products is \( \theta \), \( \theta \in [0,1] \), and the willingness to pay for remanufactured products is \( \alpha \theta \). \( \alpha \in [0,1] \) is the degree of consumer preference for remanufactured products. Due to consumers’ concerns about the quality of remanufactured products, their recognition is low. In order to improve consumers’ recognition of remanufactured products, retailers will assume corresponding corporate social responsibility for the sales of remanufactured products, and the corporate social responsibility input conducted is \( e_r \). The reference\(^8\) points out that when the CSR of the retailer's investment level of remanufactured product is \( e_r \), the willingness of consumers to pay for remanufactured product will increase by \( \beta \alpha^2 e_r \), \( \beta > 0 \) refers to the level of consumer sensitivity to corporate social responsibility. Consumers’ willingness to pay for recycled goods increased to \( \alpha \theta + \beta \alpha^2 e_r \).

At this point, the utility functions of consumers to purchase the unit of new products and remanufactured products are \( u_n = \theta - p_n, u_r = \alpha \theta - p_r + \beta \alpha^2 e_r \). The demand for new and remanufactured products is \( D_n = Q - \frac{p_n - p_r + \beta \alpha^2 e_r}{1 - \alpha}, D_r = \frac{\alpha p_n - p_r + \beta \alpha^2 e_r}{\alpha (1 - \alpha)} \). When retailers make
CSR investments, the additional costs to be borne are $C_\epsilon = \frac{1}{2} \mu \epsilon^2$. $\mu$ indicates the cost scale parameter of the retailer's CSR.

In order to ensure that recycling is economically viable for manufacturers, $B \leq C_n - C_r$. That is, $\sigma \leq C_\epsilon$, manufacturer will remanufacturing the production.

3 MODEL FORMULATION

3.1 Centralized Pricing Decision Model

Considering the connection between the upstream and downstream enterprises of the closed-loop supply chain, the income function of the whole closed-loop supply chain system considered by the system is:

$$\Pi^C = (p_n - C_n)D_n + (p_r - C_r)D_r - C(\epsilon) \tag{1}$$

The optimal solution under the centralized decision model can be derived by the inverse recursive method, as shown in Table 2.

3.2 Decentralized Pricing Decision Model

Under the decentralized decision-making mode, the profits of all participants in the closed-loop supply chain are:

$$\Pi^D_m = (\sigma_n - C_n)D_n + (\sigma_r - C_r)D_r \tag{2}$$

$$\Pi^D_r = (p_n - \sigma_n)D_n + (p_r - \sigma_r)D_r - C(\epsilon) \tag{3}$$

The equilibrium optimal solution is shown in Table 2. When $\mu > \frac{a^2 b^2}{2(1-\sigma)}$, it can be decided that $\Pi^D_n$ is a strictly concave function with respect to $p_n, p_r, \epsilon$, and $\Pi^D_r$ is a strictly concave function with respect to $\sigma_n, \sigma_r$, and there exists a unique optimal solution.

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Optimal decision of each participating system in closed-loop supply chain.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_n^*$</td>
<td>$\frac{Q + C_n}{2}$</td>
</tr>
<tr>
<td>$\sigma_r^*$</td>
<td>$\frac{aQ + C_n - C_\epsilon q}{2}$</td>
</tr>
<tr>
<td>$p_n^*$</td>
<td>$\frac{3Q + C_n}{4}$</td>
</tr>
</tbody>
</table>
4 MODEL ANALYSIS

Conclusion 1: Compared with decentralized decision-making, the selling price of new products and remanufactured products under centralized decision-making is lower, the market demand for new products and remanufactured products is higher, and the profit of the whole supply chain is higher, promoting the maximization of overall benefits and achieving win-win cooperation for both sides of the supply chain.

Proof: $p_{n}^{*} < p_{n}^{**}, p_{r}^{*} < p_{r}^{**}, D_{n}^{*} > D_{n}^{**}, D_{r}^{*} > D_{r}^{**}, \Pi_{n}^{**} > \Pi_{n}^{*}, \Pi_{r}^{**} > \Pi_{r}^{*}$.

Conclusion 2: The wholesale and retail prices of recycled products and the market demand for new products are negatively correlated with the quality of recycled products, while the social responsibility input of retailers, the market demand for remanufactured products, and the income
of retailers and manufacturers are positively correlated with the quality of recycled products.

Proof: \( \frac{\partial m^*_r}{\partial q} < 0, \frac{\partial p^*_r}{\partial q} < 0, \frac{\partial c^*_r}{\partial q} > 0, \frac{\partial d^*_m}{\partial q} < 0, \frac{\partial d^*_c}{\partial q} > 0, \)

\( \frac{\partial n^*_m}{\partial q} > 0, \frac{\partial n^*_c}{\partial q} > 0. \)

Conclusion 3: Wholesale price of remanufactured goods, retail price, CSR input of retailers, market demand for remanufactured goods and remanufactured goods preference are positively correlated, while market demand for new products and remanufacturing preference are negatively correlated.

Proof: \( \frac{\partial m^*_r}{\partial a} > 0, \frac{\partial p^*_r}{\partial a} > 0, \frac{\partial c^*_r}{\partial a} > 0, \frac{\partial d^*_m}{\partial a} < 0, \frac{\partial d^*_c}{\partial a} > 0. \)

Conclusion 4: The wholesale price of new and remanufactured products and the retail price of new products are not related to consumer sensitivity to CSR, while the retail price of remanufactured products, retailer CSR input, market demand for remanufactured products, and retailer and manufacturer revenue are all positively related to consumer sensitivity to CSR, and market demand for new products is negatively related to consumer sensitivity to CSR.

Proof: \( \frac{\partial p^*_r}{\partial \beta} > 0, \frac{\partial c^*_r}{\partial \beta} > 0, \frac{\partial d^*_m}{\partial \beta} > 0, \frac{\partial d^*_c}{\partial \beta} < 0, \frac{\partial n^*_m}{\partial \beta} > 0, \)

\( \frac{\partial n^*_c}{\partial \beta} > 0. \)

5 NUMERICAL STUDIES

In order to verify the rationality and validity of the above findings, numerical analysis is conducted in this section, focusing on the impact of the quality of recycled products and consumer preferences for recycled products as well as the sensitivity of CSR on the equilibrium results of the supply chain. According to the relevant literature and the constraint assumptions \( \alpha = 0.5, \beta = 3, C_r = 150, Q = 1000, C_e = 120, \mu = 10 \), numerical simulations were conducted using Maple software to analyze the trend of the effect of recycled product quality on the retail price of recycled products, CSR inputs, product demand and supply chain revenue of the firm under the two decision models. \( q \in [0.5, 1] \) in order to ensure that the variables were meaningful, as shown below.
According to Figure 2, with the increase in the quality level of recycled products, consumers' enthusiasm to participate in recycling activities increases with the incentive of recycled product subsidies, saving the remanufacturing cost of the remanufacturing process, thus promoting the reduction of the price of remanufactured products and increasing the market share of remanufactured products, while retailers' CSR investment in recycled products, i.e. charity level, also increases, which has a positive impact on increasing the overall profit of the supply chain without affecting the price of new products.

When studying the influence trend of consumers' preference for the remanufactured product on the retail price of the remanufactured product, CSR input, product demand and supply chain income, in order to ensure that the variables have a meaningful $\alpha \in (0.6,0.8)$, as shown in the figure below.
As can be seen from Figure 3, under the two decision modes, with the increase of consumers' preference for the remanufactured product, the retail price of the remanufactured product and the CSR input of retailers will increase, and consumers' demand for the remanufactured product will also increase gradually. When consumers have the same preference for recycled goods, the retail price of centralized decision-making is lower, and retailers can gain a larger market share through the lower retail price, which has a positive impact on promoting the sales of remanufactured goods and the overall income of the supply chain.

In analyzing the trends of the effects of retailers' CSR input costs and consumer sensitivity on firms' retail prices of remanufactured products, CSR inputs, product demand, and supply chain revenue under the two decision models, in order to ensure that the variables are meaningful, $\beta \in [0.5, 1.5], \mu \in [2, 5]$ As shown in the figure below:
According to Figure 4, the retail price of recycled goods, the level of CSR input by retailers and the profit of the overall supply chain increase with the increase of consumer sensitivity, then the cost of CSR input will also increase accordingly, because consumers are willing to actively bear the CSR input by retailers, so the price increase does not affect the demand of recycled goods, while the increase of cost in the process of CSR input will reduce the corresponding enterprise. The increase in the cost of CSR input process will reduce the corresponding CSR input of enterprises, resulting in a corresponding decrease in the market demand for recycled products, thus leading to a reduction in the overall profit of the closed-loop supply chain. Therefore, enterprises should focus on the innovation ability of member companies, reduce the cost consumption in the process of CSR input, and improve the efficiency of CSR input level, so that they can actively perform CSR behavior and improve their own profit at the same time.

From the simulation diagram and the above analysis, we can see that the centralized decision pricing model is better than the decentralized decision model.

6 CONCLUSION

Based on the consumer preferences of different products, this paper studies the impact of different recovery quality and retailers' CSR investment on the pricing of remanufactured products and the revenue of remanufactured supply chain. By building centralized and decentralized decision-making models, it provides reference for enterprises to make decisions. Through the analysis, the following conclusions are drawn:

Firstly, enterprises should pay attention to the efficiency of CSR inputs to reduce input costs while increasing consumer attention; secondly, the government and enterprises can conduct publicity to improve consumer awareness of recycled products and enhance consumer recognition and preference for recycled products; thirdly, since there is a relationship between retailers' social donations and the quality of recycled products, enterprises and the government
can make efforts to introduce relevant subsidies and welfare policies to encourage consumers to provide high-quality recycled goods, thereby improving retailers' CSR donation levels and closed-loop supply chain profits, forming a virtuous cycle.

In the research process, the comparison found that the centralized decision model is still the optimal decision, so the coordination mechanism of the supply chain can be further discussed in the subsequent research; at the same time, we only focused on the forward pricing study of recycled goods in the study, and the next step can be based on this to consider the optimal pricing study of recycled goods in the reverse recycling process.

REFERENCES

Statistical Measurement of Spatial and Temporal Differences in Regional Economic Growth in China

Haiyan Wang¹, Sirong Shang¹*, Zhanni Huang¹

*3380511836@qq.com
School of Statistics and Mathematics, Yunnan University of Finance and Economics, No. 295 Longquan Road, Kunming, China¹

Abstract: This article is devoted to exploring the spatial and temporal equilibrium of China’s regional economic growth, constructing a regional economic measurement system based on the requirements of the new normal of economic development, and using factor analysis to statistically measure the economic growth levels of 31 provinces and municipalities in mainland China from 2013-2019. The results show that China’s regional economic growth has different characteristics at various levels of economic development, social life, industrial production and resource utilization, among which economic development and social life are the key factors affecting the balance of China’s economic growth. In addition, the economic growth level of each province and city has different characteristics in different dimensions, and the overall trend is “decreasing from east to west”. Finally, we hope to reveal the development trend and spatial differences of China’s regional economic growth under the new normal through the statistical measurement of various provinces and cities and various indicators, and to contribute Chinese solutions and experiences to achieve sustainable and balanced economic development.

Keywords: Regional Economy, Factor Analysis, Economic Growth, Statistical Measures.

1 INTRODUCTION

Since the reform and opening up of China, the continuous optimization of China’s economic structure has enabled the economy to maintain a long-term trend of steady growth, and in 2010, China’s average annual GDP growth rate surpassed that of Japan to become the world’s second largest economy after the United States. However, the imbalance between China’s regional economic development and the incompatibility between economic growth and resource conservation have become two major potential constraints to the balanced growth of China’s economy. With the development of China’s economy and society, excessive resource consumption and negative economic growth have become inevitable core problems. The increasing dependence of economic growth on investment demand, the inefficiency of input and output of various factors within the economy, and the unreasonable allocation of resources are a series of economic risks that will hinder the coordinated development of the regional economy.

As economic growth is the key issue of regional economics research, there should be scientific theoretical guidance to achieve high-quality economic growth and coordinated economic development among regions in China accordingly, so as to better reveal the essential requirements and social laws of high-quality economic development. Based on the requirement
of high-quality and coordinated regional economic development, narrowing the differences in economic development among regions and dealing with the relationship between economic growth and ecological sustainability have become the top priority nowadays. How to better measure the level of regional economic development in China and analyze the factors influencing the differences in economic growth is the focus of this paper.

The report of the 19th Party Congress emphasizes that China’s economy has transformed from high-speed growth to high-quality development, which means that high-quality economic growth model has become the core lever of China’s economic development in the new era, and also indicates that China’s economic construction in the future should adhere to high-quality and high-standard as the measure of development. However, at this stage, the research on the economic growth status of each region in China is still mainly at the level of evaluation of high-quality economic development, and the research on the influencing factors of regional economic differentiation is also slightly insufficient.

2 LITERATURE REVIEW

China’s high-quality economic development and economic synergistic growth fit the inherent requirements of the new development concept. From the perspective of economics, high-quality economic development is committed to meeting the people’s growing value and spiritual pursuit of a better life dimension, reflecting the essential characteristics of economic development; while economic synergistic growth reflects the docking of products and services based on the perspective of supply-side structural reform to enhance the intrinsic value of China’s economic policies. The organic combination of the two is the inherent embodiment of achieving common prosperity and synergistic economic growth, which is in line with the inherent requirements of the new development concept. From the current research literature, academic research on the issue of high-quality development and economic growth differentials in China’s economy is still in the exploratory stage.

Yachen Shen and Rui Wu (2022) constructed a regional digital economy development measurement index system and used the entropy value method to measure the level of regional digital economy development from two dimensions of industrial digitization and digital industry, and found that investment efforts, government regulation mechanism, transportation infrastructure construction and education level are the significant reasons for the differences between the digital economy and the regional economic quality development. Ju Pan and Dan Yang (2019) used quantile regression to confirm the association between the differences in capital, human, technology and institutional economic factors and the regional economic development gap, and found that the differences in capital and human factor inputs are the reasons for the significant differences in regional economic development in China. Ying Li (2020) illustrated the differences in economic growth levels among Chinese regions from the perspective of geographic economics, and analyzed the agglomeration effect and spatial heterogeneity of each economic factor. Min Wei and Shuhao Li (2018) confirmed the existence of a close correlation between high-quality economic development and economic growth, and analyzed the changes in spatial and temporal differences of regional economic development from the perspective of economic growth, further analyzing the coordination of China’s economic development level and economic policies, responding to the conceptual requirements
of China’s economic development in the new era, and also summarizing and sublimating the theory of economic growth measurement. Baoping Ren et al.(2015) indicate that there is some correlation between the overall and regional economies in China, and there are temporal and spatial differences in the quality of economic growth between the overall and regional economies. Jie Wei et al.(2021) conducted a deeper investigation and consideration on the issue of regional economic growth efficiency and development imbalance. Chunhui Gan et al.(2011) constructed the economic growth index from four dimensions: economic structure, growth status, changes in economic growth welfare and resource sustainability, and then measured the regional economic development. Kun Li(2019) measured China’s economic development from the perspectives of economic growth dynamics, structural differences, social welfare and resource environment. Wei Wang(2020) selected the indicators of the economy of scale, industrial structure and ecological coordination to construct the measurement system, and used factor analysis and principal component analysis to measure the regional economic development level. Ying Xu et al.(2004) constructed the system of high-quality economic development from three dimensions, including social economy, industrial sector and market factors, and focused on the differences in economic development among Chinese regions. The measurement of regional economic development level differences in the literature is also the research direction of this paper.

In general, the research on the development of China’s regional economies and the differences in economic growth should be focused on. Therefore, this paper draws on the measurement ideas of the existing literature and uses factor analysis to study and analyze the differences in regional economic development and economic growth in China and the influencing factors, and determines the association between the indicators of each economic factor and the differences in economic growth in the context of the regional economy through the weight values of indicators, which enriches the research theory and practical reference on the measurement of the regional economy.

3 DATA SELECTION AND THEORETICAL MODEL

3.1 Data Selection

Due to the influence of many comprehensive factors, regional differences in China’s economic growth appear. Therefore, regional economic evaluation indicators should be selected from multiple perspectives, and an economic measurement and evaluation system should be constructed to analyze the reasons for the differentiation of economic growth among regions by measuring the degree of influence of their indicators, and to identify the important influencing factors. To ensure the scientificity and rigor of the data, this paper reviews the relevant data information of the China Statistical Yearbook and Environmental Statistical Yearbook, and on the basis of the research results of Liu Guobin, Song Jinze, Ren Baoping and other scholars, selects the representative economic development indicators in 31 provinces and cities in inland China from 2013 to 2019, and finally constructs the indicator evaluation system based on the principles of representativeness, comprehensiveness and comparability as well as China’s philosophy about regional economic synergistic high-quality development under the new normal, the complex regional economic synergistic development and endogenous growth concept is decomposed into four major dimensions of economic development, social life,
industrial production and resource utilization, and further subdivided into 14 indicators, as shown in Table 1.

<table>
<thead>
<tr>
<th>System level</th>
<th>Indicator Level</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Economic Development</td>
<td>GDP</td>
<td>$X_1$</td>
</tr>
<tr>
<td></td>
<td>GDP per capita</td>
<td>$X_2$</td>
</tr>
<tr>
<td></td>
<td>GDP Growth Rate</td>
<td>$X_3$</td>
</tr>
<tr>
<td>Social life</td>
<td>Natural Population Growth Rate</td>
<td>$X_4$</td>
</tr>
<tr>
<td></td>
<td>Urban Registered Unemployment Rate</td>
<td>$X_5$</td>
</tr>
<tr>
<td></td>
<td>Per capita consumption expenditure of urban households</td>
<td>$X_6$</td>
</tr>
<tr>
<td>Industrial Production</td>
<td>Chemical Oxygen Demand Emission</td>
<td>$X_7$</td>
</tr>
<tr>
<td></td>
<td>Industrial Dust Emission</td>
<td>$X_8$</td>
</tr>
<tr>
<td></td>
<td>Industrial wastewater</td>
<td>$X_9$</td>
</tr>
<tr>
<td></td>
<td>Industrial fixed waste</td>
<td>$X_{10}$</td>
</tr>
<tr>
<td></td>
<td>Industrial sulfur dioxide emissions</td>
<td>$X_{11}$</td>
</tr>
<tr>
<td>Resource Utilization</td>
<td>Energy consumption per unit GDP</td>
<td>$X_{12}$</td>
</tr>
<tr>
<td></td>
<td>Water consumption per unit GDP</td>
<td>$X_{13}$</td>
</tr>
<tr>
<td></td>
<td>Fixed asset investment per unit of GDP</td>
<td>$X_{14}$</td>
</tr>
</tbody>
</table>

The statistical distribution of the number of indicators is shown in Figure 1 below.
3.2 Measurement Models and Methods

Thurstone first introduced the concept of factor analysis method in 1931, which was initially used for simple statistical analysis of data, but with the in-depth exploration and development of the discipline of computational science, factor analysis has been widely used in various statistical disciplines. Factor analysis aims at describing statistics and interpreting the components of the original observations by summing linear and special factors with a minimum number of unpredictable factors. The greatest advantage of the factor analysis method is the practicality of its dimensionality reduction idea, which can convert multiple indicators into a small number of uncorrelated indicators, thus transforming a complex analysis into a simple multivariate analysis method, allowing the data to reflect the relational nature between them by itself. Therefore, this paper adopts the factor analysis method to measure the economic development level of each region in China, so as to ensure the rationality and scientficity of the empirical quantitative analysis as much as possible from the root.

Let $X = (x_1, x_2, \cdots, x_p)^T$ be the observable variables of the economic development level measure in this paper, and due to the different unit and dimensionless process among the variables, the standardized variables $\bar{x_i}$ are obtained by using the polarization method of equation (1), $K = (k_1, k_2, \cdots, k_q)^T$ as factor vector, $M = T_{ij}(i = 1, 2, \cdots, p; j = 1, 2, \cdots, q)$ is

![Figure 1: Statistical results of the number of indicators](image-url)
the factor loading matrix, and \( \delta = (\delta_1, \delta_2, \cdots, \delta_p)^T \) is the special factor. The factor analysis model equation (2) and equation (3) are obtained as follows.

\[
\bar{x}_i = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{1}
\]

\[
\bar{x}_1 = T_{11}k_1 + T_{12}k_2 + \cdots + T_{1q}k_q + \delta_1
\]

\[
\bar{x}_2 = T_{21}k_1 + T_{22}k_2 + \cdots + T_{2q}k_q + \delta_2
\]

\[\vdots\]

\[
\bar{x}_p = T_{p1}k_1 + T_{p2}k_2 + \cdots + T_{pq}k_q + \delta_p \tag{2}
\]

\[X = KM + \delta = KT_{ij} + \delta \tag{3}\]

The sum of the squared terms of the elements in the \( i \)th row of the loading matrix \( M \) is denoted as \( \sum T_{ij}^2 \), \( (i = 1, 2, \cdots, p) \), then the co-dependence of the indicator variable \( X \) is the co-dependence of the common factor \( K \) on the variable \( X \), which can reflect the dependence of all the co-dependent factors on the indicator variable \( X \). The larger the value of \( \sum T_{ij}^2 \), the greater the dependence of the \( i \)th indicator on each of the common factors \( K \), and vice versa. Similarly, the sum of the squared terms of the elements of the \( j \)th column in the loading matrix \( M \) is denoted as \( \sum T_{ij}^2 \), \( (j = 1, 2, \cdots, q) \) \( . \) \( \sum T_{ij}^2 \) indicates the variance contribution of the public factor \( K \) to the variable \( X \). The larger the value, the higher the contribution of the \( j \)th indicator to each public factor \( K \).

After the dependence and contribution calculation of the factor model, the factor loadings need to be rotated.

Let \( U \) be an orthogonal matrix, and the rotated factor loading matrix \( L = l_{ij} \) \( (i = 1, 2, \cdots, p; j = 1, 2, \cdots, q) \) is obtained by equation (4). Further, a new factor analysis model is constructed as shown in equations (5) and (6) below.

\[L = MU \tag{4}\]

\[
\bar{x}_1 = l_{11}k_1 + l_{12}k_2 + \cdots + l_{1q}k_q + \delta_1
\]

\[
\bar{x}_2 = l_{21}k_1 + l_{22}k_2 + \cdots + l_{2q}k_q + \delta_2
\]

\[\vdots\]

\[
\bar{x}_p = l_{p1}k_1 + l_{p2}k_2 + \cdots + l_{pq}k_q + \delta_p \tag{5}
\]

\[X = KMU + \delta = Kl_{ij} + \delta \tag{6}\]

Finally, each factor is expressed as a linear combination equation (7) based on the original variables, from which the composite score of each factor is calculated.

\[
E_j = l_{ij}x_1 + l_{ij}x_2 + \cdots + l_{ij}x_p \tag{7}\]
4 EMPIRICAL ANALYSIS

4.1 Descriptive Statistics

This paper refers to the relevant statistical data of the China Statistical Yearbook and Environmental Statistical Yearbook, selects the economic indicators of 31 provinces and municipalities in China from 2013 to 2019, and constructs the measurement and evaluation system. In addition, for the consideration of data authenticity, validity and predictability, the article standardizes the selected data, and some missing data are complemented by Lagrangian interpolation method, and finally 217 valid samples are obtained, corresponding to the principle method of interpolation as follows.

By constructing a plane containing \( n \) points, we can obtain a polynomial equation that passes through \( n \) points in the plane and contains \( n - 1 \) terms to the \( n \)th power.

\[
y = \beta_0 + \beta_1 x + \beta_2 x^2 + \cdots + \beta_{n-1} x^{n-1}
\]  

The coordinates of these \( n \) points in the plane \((x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\) are then substituted into the polynomial (8) to obtain a system of \( n \) equations.

\[
\begin{align*}
y_1 &= \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 + \cdots + \beta_{n-1} x_1^{n-1} \\
y_2 &= \beta_0 + \beta_1 x_2 + \beta_2 x_2^2 + \cdots + \beta_{n-1} x_2^{n-1} \\
&\vdots \\
y_n &= \beta_0 + \beta_1 x_n + \beta_2 x_n^2 + \cdots + \beta_{n-1} x_n^{n-1}
\end{align*}
\]  

which leads to the solution of the Lagrangian system of equations.

\[
L(x) = y_i \frac{(x-x_2)(x-x_3)\cdots(x-x_n)}{(x_1-x_2)(x_1-x_3)\cdots(x_1-x_n)} + y_2 \frac{(x-x_1)(x-x_3)\cdots(x-x_n)}{(x_2-x_1)(x_2-x_3)\cdots(x_2-x_n)} + \cdots + y_n \frac{(x-x_1)(x-x_2)\cdots(x-x_{n-1})}{(x_n-x_1)(x_n-x_2)\cdots(x_n-x_{n-1})}
\]  

Based on equation (10) above, the approximate value \( L(x) \) of the missing value is obtained by substituting the point \( x \) corresponding to the missing function value into the interpolation polynomial.

The results of descriptive statistics were further analyzed after interpolation of the data, as shown in Table 2 below.
As can be seen from Table 2, the mean value of GDP is 2,564.157 5 billion yuan and the GDP per capita is 57,652.806 yuan, indicating that China’s economic development in this period showed a good trend and adapted to the requirements of high-quality sustainable economic development. However, the standard deviations corresponding to both indicators are large, which reflects that while China’s economy continues to grow, there are still large differences in economic development among regions. In addition, the average GDP growth rate is 7.796%, which is greater than 3%, indicating that China’s economic development in this period has a certain scale and a positive development trend, and the living standard of urban and rural residents has been improved.
As can be seen from the results of Figure 2, only the data corresponding to GDP growth rate \((X_3)\), natural population growth rate \((X_8)\), urban registered unemployment rate \((X_5)\) and water consumption per unit of GDP \((X_9)\) as a whole approximately obey a normal distribution, while the rest of the indicators show skewed distribution. It indicates that these four data indicators have good symmetry and stability. It indicates that the GDP level and the natural population growth rate of Chinese provinces and cities maintain a relatively stable growth rate during the period 2013-2019. In addition, the data distribution of the urban registered unemployment rate in each province and city between these seven years is more concentrated than the rest of the indicators, indicating that the urban unemployment rate in each region of China has maintained a stable trend during this period. Similarly, the data distribution of the unit GDP water consumption indicator also indicates that the water consumption of the regional GDP in each province and city of China is basically stable in general. The distribution of the coefficient of variation shows that the \(X_{14}\) indicators are more discrete in their data distribution, which also indicates that there are significant spatial and temporal differences in the amount of fixed asset investment in China’s regional GDP.

### 4.2 KMO Test and Bartlett’s Sphericity Test

The KMO test is based on the principal component analysis to select the factors with representative significance and significant correlation from the original many factors, and
compare them systematically with the original variables to analyze the correlation degree coefficient and bias correlation coefficient between them. If the KMO value tends to 1, it means that the correlation between the factor and the original variable is strong, and vice versa. The selection of representative factors is based on whether the KMO test value is greater than 0.7. If the test value is greater than 0.7, it means that the selected factors are more suitable for the measurement model of this paper. The expression of the KMO test is constructed as follows.

\[
KMO = \frac{\sum \sum_i \sum_j r_{ij}^2}{\sum i \not= j r_{ij}^2 + \sum \sum_{i \not= j, 1, 2, \ldots, K} r_{ij}^2}
\]  

Bartlett’s spherical test presupposes the construction of the correlation coefficient matrix of the factors, and specifies \( H_0 \): the correlation coefficient matrix is a unit array as the original hypothesis condition. Further based on the determinant of the correlation coefficient matrix to calculate its statistics, after testing if the Bartlett’s approximate chi-square value is large and the corresponding central value of the probability distribution is below the significant level, then in principle \( H_0 \) is rejected and there is an association between the original variables and the factors, which can be used for factor analysis. Conversely, \( H_0 \) cannot be rejected.

The Kaiser test principle is used to determine whether there is a common factor among the variable indicators, and the limits of the KMO values are also judged to analyze the correlation between the variable indicators. The test results are shown in Table 3.

<table>
<thead>
<tr>
<th>Test Category</th>
<th>Range of values</th>
<th>Factor adaptation</th>
<th>Kaiser-Meyer-Olkin</th>
<th>Bartlett’s approximate cardinality</th>
<th>df</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>KMO</td>
<td>&gt;0.9</td>
<td>Very suitable</td>
<td>0.805</td>
<td></td>
<td>91.000</td>
<td>0.000***</td>
</tr>
<tr>
<td></td>
<td>0.8~0.9</td>
<td>Very suitable</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.7~0.8</td>
<td>Suitable</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.6~0.7</td>
<td>Barely suitable</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5~0.6</td>
<td>Not very suitable</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>&lt;0.5</td>
<td>Not suitable</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Based on the requirement of the adequacy of the Kaiser-Meyer-Olkin sample, the numerical magnitude of the correlation coefficient and the bias correlation coefficient of the two factor variables need to be compared, which in turn leads to an important measure of the strength of the correlation between the representative factor and the original variable. After testing, the KMO value of this paper’s model is 0.805 (0.8, 0.9); the Bartlett’s approximate chi-square value is 2350.713, which is a large value, and the test significance level result is less than 0.05, and the original hypothesis is rejected in principle. It indicates that the data in this paper are suitable for the factor analysis model.

### 4.3 Principal Component Analysis

The principal component analysis is mainly used for the compression and interpretation of factor information and the amplification of principal component information of the original data. Its
analysis mainly includes 3 elements, namely, characteristic root, variance contribution rate and cumulative contribution rate. The explanatory power of the principal components is judged by the characteristic root index, if the characteristic root is less than 1, it means that the explanatory power of the selected principal component factors is slightly lower than that of the original variables, and vice versa; the variance contribution rate has a positive correlation with the characteristic root, the higher the variance contribution rate, the stronger the ability to extract explanations for the information; the cumulative contribution rate indicates the amount of information extracted by the first $n$ principal component factors, which can also be interpreted as the principal component The cumulative contribution rate indicates the cumulative amount of information extracted by the first $n$ principal component factors, which can also be interpreted as the combined contribution influence of the principal component factors to the overall sample. It is usually considered that the total contribution of the first $n$ principal component factors exceeds 80%, which can indicate that these first $n$ principal component factors can explain the overall information of the sample.

Table 4: Principal component feature roots

<table>
<thead>
<tr>
<th>Variables</th>
<th>Initial Eigenvalue</th>
<th>Extraction of the sum of squares of loads</th>
<th>Sum of squared rotating loads</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Percentag e of variance</td>
<td>Accumulati on</td>
</tr>
<tr>
<td>$X_3$</td>
<td>1.844</td>
<td>13.175</td>
<td>64.253</td>
</tr>
<tr>
<td>$X_4$</td>
<td>1.198</td>
<td>8.559</td>
<td>72.813</td>
</tr>
<tr>
<td>$X_5$</td>
<td>1.018</td>
<td>7.273</td>
<td>80.086</td>
</tr>
<tr>
<td>$X_6$</td>
<td>0.741</td>
<td>5.290</td>
<td>85.376</td>
</tr>
<tr>
<td>$X_7$</td>
<td>0.610</td>
<td>4.361</td>
<td>89.737</td>
</tr>
<tr>
<td>$X_8$</td>
<td>0.513</td>
<td>3.661</td>
<td>93.398</td>
</tr>
<tr>
<td>$X_9$</td>
<td>0.314</td>
<td>2.243</td>
<td>95.641</td>
</tr>
<tr>
<td>$X_{10}$</td>
<td>0.224</td>
<td>1.599</td>
<td>97.241</td>
</tr>
<tr>
<td>$X_{11}$</td>
<td>0.175</td>
<td>1.253</td>
<td>98.494</td>
</tr>
<tr>
<td>$X_{12}$</td>
<td>0.128</td>
<td>0.912</td>
<td>99.406</td>
</tr>
<tr>
<td>$X_{13}$</td>
<td>0.060</td>
<td>0.431</td>
<td>99.837</td>
</tr>
<tr>
<td>$X_{14}$</td>
<td>0.023</td>
<td>0.163</td>
<td>100.000</td>
</tr>
</tbody>
</table>

As can be seen from Table 4, the first five principal component factors correspond to eigenroots greater than 1, and the cumulative variance contribution rate is 80.086% > 80%, indicating that the first five indicators in the data selected for this paper have strong explanatory power of information for the overall sample, and are also the main factors affecting the differences in regional economic growth, namely: GDP, GDP per capita, GDP growth rate, natural population growth rate and urban registered unemployment rate. The characteristic root fragmentation plots corresponding to the 14 indicators tested are as follows.
As can be seen in Figure 3, the eigenvalue curves within the interval of the first five variables are steeper, indicating that the first five factor variables explain the research problem of this paper to a higher degree; when the number of variable factors reaches six, their eigenvalue curves tend to level off, so the first five variables are used as the extracted factors. Further, principal component analysis was applied to extract the variable factors to construct the factor loading matrix, and the quantitative relationship between each factor and the original variables was analyzed by observing the coefficients of the matrix. The rotated component matrix was obtained after quantifying and eliminating the unreasonable variables in the matrix, so that the correspondence between all data variables and the selected factors met the expected effect. The results are shown in Tables 5 and 6.

**Table 5: Component matrix**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Component Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Factor 1</td>
</tr>
<tr>
<td>$X_1$</td>
<td>-0.046</td>
</tr>
<tr>
<td>$X_2$</td>
<td>-0.670</td>
</tr>
<tr>
<td>$X_3$</td>
<td>-0.164</td>
</tr>
<tr>
<td>$X_4$</td>
<td>0.045</td>
</tr>
<tr>
<td>$X_5$</td>
<td>0.496</td>
</tr>
<tr>
<td>$X_6$</td>
<td>-0.764</td>
</tr>
<tr>
<td>$X_7$</td>
<td>0.605</td>
</tr>
<tr>
<td>$X_8$</td>
<td>0.749</td>
</tr>
<tr>
<td>$X_9$</td>
<td>0.156</td>
</tr>
<tr>
<td>$X_{10}$</td>
<td>0.579</td>
</tr>
<tr>
<td>$X_{11}$</td>
<td>0.794</td>
</tr>
<tr>
<td>$X_{12}$</td>
<td>0.542</td>
</tr>
<tr>
<td>$X_{13}$</td>
<td>0.609</td>
</tr>
<tr>
<td>$X_{14}$</td>
<td>-0.496</td>
</tr>
</tbody>
</table>
Table 6: Component matrix after rotation

<table>
<thead>
<tr>
<th>Variables</th>
<th>Component Factors</th>
<th>Common factor variance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Factor 1</td>
<td>Factor 2</td>
</tr>
<tr>
<td>$X_{1,s}$</td>
<td>0.940</td>
<td>0.160</td>
</tr>
<tr>
<td>$X_{2,s}$</td>
<td>0.503</td>
<td>-0.202</td>
</tr>
<tr>
<td>$X_{3,s}$</td>
<td>0.201</td>
<td>-0.185</td>
</tr>
<tr>
<td>$X_{4,s}$</td>
<td>-0.154</td>
<td>-0.024</td>
</tr>
<tr>
<td>$X_{5,s}$</td>
<td>-0.074</td>
<td>0.149</td>
</tr>
<tr>
<td>$X_{6,s}$</td>
<td>0.394</td>
<td>-0.302</td>
</tr>
<tr>
<td>$X_{7,s}$</td>
<td>0.511</td>
<td>0.294</td>
</tr>
<tr>
<td>$X_{8,s}$</td>
<td>0.129</td>
<td>0.862</td>
</tr>
<tr>
<td>$X_{9,s}$</td>
<td>0.910</td>
<td>0.133</td>
</tr>
<tr>
<td>$X_{10,s}$</td>
<td>0.072</td>
<td>0.881</td>
</tr>
<tr>
<td>$X_{11,s}$</td>
<td>0.165</td>
<td>0.674</td>
</tr>
<tr>
<td>$X_{12,s}$</td>
<td>-0.569</td>
<td>0.550</td>
</tr>
<tr>
<td>$X_{13,s}$</td>
<td>-0.033</td>
<td>0.011</td>
</tr>
<tr>
<td>$X_{14,s}$</td>
<td>0.000</td>
<td>-0.097</td>
</tr>
</tbody>
</table>

From the rotated factor component matrix in Table 6 and the rotated factor loadings heat map in Figure 4, it can be seen that factor 1 has high loadings on variables $X_{1,s}$, $X_{2,s}$ and $X_{9,s}$, indicating that factor 1 is an economic factor; similarly, since factor 2 has high loadings on variables $X_{6,s}$ and $X_{10,s}$, factor 2 is defined as an industrial factor; factor 3 has high loading on variable $X_{13,s}$, defined as the energy consumption factor; factor 4 has a higher loading on variable $X_{14,s}$, defined as the investment factor; and factor 5 has a higher loading on variables $X_{5,s}$ and $X_{6,s}$, defined as the social factor.

Figure 4: Heat map of rotational factor loadings
4.4 Overall Score Assessment

In order to further investigate the factors affecting the balanced development of China’s regional economies and causing differences in economic growth, this paper uses factor analysis to calculate the combined score coefficients of each factor, as shown in Table 7.

Table 7: Corresponding feature vectors of principal components

<table>
<thead>
<tr>
<th>Component Matrix</th>
<th>Main component 1</th>
<th>Main component 2</th>
<th>Main component 3</th>
<th>Main component 4</th>
<th>Main component 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variables</td>
<td>( k_1 )</td>
<td>( k_2 )</td>
<td>( k_3 )</td>
<td>( k_4 )</td>
<td>( k_5 )</td>
</tr>
<tr>
<td>( X_1 )</td>
<td>0.227</td>
<td>0.053</td>
<td>-0.042</td>
<td>0.038</td>
<td>-0.002</td>
</tr>
<tr>
<td>( X_2 )</td>
<td>0.122</td>
<td>-0.067</td>
<td>-0.287</td>
<td>0.317</td>
<td>0.377</td>
</tr>
<tr>
<td>( X_3 )</td>
<td>0.049</td>
<td>-0.061</td>
<td>-0.022</td>
<td>0.014</td>
<td>-0.656</td>
</tr>
<tr>
<td>( X_4 )</td>
<td>-0.037</td>
<td>-0.008</td>
<td>0.033</td>
<td>0.001</td>
<td>-0.823</td>
</tr>
<tr>
<td>( X_5 )</td>
<td>-0.018</td>
<td>0.05</td>
<td>0.096</td>
<td>-0.579</td>
<td>0.446</td>
</tr>
<tr>
<td>( X_6 )</td>
<td>0.095</td>
<td>-0.1</td>
<td>-0.307</td>
<td>0.35</td>
<td>0.364</td>
</tr>
<tr>
<td>( X_7 )</td>
<td>0.124</td>
<td>0.098</td>
<td>0.382</td>
<td>-0.031</td>
<td>0.029</td>
</tr>
<tr>
<td>( X_8 )</td>
<td>0.031</td>
<td>0.286</td>
<td>0.14</td>
<td>-0.071</td>
<td>0.132</td>
</tr>
<tr>
<td>( X_9 )</td>
<td>0.22</td>
<td>0.044</td>
<td>0.129</td>
<td>0.002</td>
<td>-0.034</td>
</tr>
<tr>
<td>( X_{10} )</td>
<td>0.018</td>
<td>0.292</td>
<td>-0.05</td>
<td>-0.139</td>
<td>0.073</td>
</tr>
<tr>
<td>( X_{11} )</td>
<td>0.04</td>
<td>0.224</td>
<td>0.308</td>
<td>-0.022</td>
<td>0.037</td>
</tr>
<tr>
<td>( X_{12} )</td>
<td>-0.137</td>
<td>0.182</td>
<td>0.089</td>
<td>-0.062</td>
<td>0.092</td>
</tr>
<tr>
<td>( X_{13} )</td>
<td>-0.008</td>
<td>0.004</td>
<td>0.471</td>
<td>-0.106</td>
<td>0.047</td>
</tr>
<tr>
<td>( X_{14} )</td>
<td>0.0</td>
<td>-0.032</td>
<td>-0.047</td>
<td>0.766</td>
<td>0.13</td>
</tr>
</tbody>
</table>

Combining the total variance of the original variables explained in Table 5 and the corresponding eigenvectors of the principal components in Table 7, the combined score function of each factor can be constructed.

\[
E_1 = 0.227X_1 + 0.122X_2 + 0.049X_3 - 0.037X_4 - 0.018X_5 + 0.095X_6 + 0.124X_7 + 0.031X_8 + 0.22X_9 + 0.018X_{10} + 0.04X_{11} - 0.137X_{12} - 0.008X_{13} + 0.0X_{14}
\]

\[
E_2 = 0.053X_1 - 0.067X_2 - 0.061X_3 - 0.008X_4 + 0.05X_5 - 0.1X_6 + 0.098X_7 + 0.286X_8 + 0.044X_9 + 0.292X_{10} + 0.224X_{11} + 0.182X_{12} + 0.004X_{13} - 0.032X_{14}
\]

\[
E_3 = -0.042X_1 - 0.287X_2 - 0.022X_3 + 0.033X_4 + 0.096X_5 - 0.307X_6 + 0.382X_7 + 0.14X_8 + 0.129X_9 - 0.05X_{10} + 0.308X_{11} + 0.089X_{12} + 0.471X_{13} - 0.047X_{14}
\]

\[
E_4 = 0.03X_1 + 0.317X_2 + 0.014X_3 + 0.001X_4 - 0.579X_5 + 0.35X_6 - 0.031X_7
\]
Finally, the composite factor score function is derived from Equation (7).

\[
E = \frac{0.202}{0.801} \times E_1 + \frac{0.186}{0.801} \times E_2 + \frac{0.17}{0.801} \times E_3 + \frac{0.122}{0.801} \times E_4 + \frac{0.121}{0.801} \times E_5
\]

(12)

Finally, the comprehensive economic development ranking of each province and city is calculated based on the comprehensive factor score function (12) formula, as shown in Table 8.

<table>
<thead>
<tr>
<th>Province</th>
<th>(E_1)</th>
<th>(E_2)</th>
<th>(E_3)</th>
<th>(E_4)</th>
<th>(E_5)</th>
<th>Overall Score((E))</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guangdong</td>
<td>2.735</td>
<td>-0.233</td>
<td>0.863</td>
<td>0.555</td>
<td>-0.812</td>
<td>0.78</td>
<td>1</td>
</tr>
<tr>
<td>Shandong</td>
<td>1.526</td>
<td>1.545</td>
<td>0.021</td>
<td>-0.006</td>
<td>-0.384</td>
<td>0.689</td>
<td>2</td>
</tr>
<tr>
<td>Jiangsu</td>
<td>2.234</td>
<td>0.301</td>
<td>-0.353</td>
<td>-0.083</td>
<td>0.342</td>
<td>0.597</td>
<td>3</td>
</tr>
<tr>
<td>Liaoning</td>
<td>-0.045</td>
<td>1.06</td>
<td>0.176</td>
<td>0.086</td>
<td>1.793</td>
<td>0.556</td>
<td>4</td>
</tr>
<tr>
<td>Hebei</td>
<td>0.073</td>
<td>2.349</td>
<td>-0.303</td>
<td>-0.211</td>
<td>-0.042</td>
<td>0.461</td>
<td>5</td>
</tr>
<tr>
<td>Beijing</td>
<td>-0.057</td>
<td>-0.653</td>
<td>-0.809</td>
<td>4.724</td>
<td>0.432</td>
<td>0.446</td>
<td>6</td>
</tr>
<tr>
<td>Henan</td>
<td>0.744</td>
<td>0.409</td>
<td>0.583</td>
<td>-0.147</td>
<td>-0.505</td>
<td>0.307</td>
<td>7</td>
</tr>
<tr>
<td>Inner Mongolia</td>
<td>-0.574</td>
<td>2.019</td>
<td>-0.977</td>
<td>-0.064</td>
<td>0.972</td>
<td>0.254</td>
<td>8</td>
</tr>
<tr>
<td>Shanxi</td>
<td>-0.962</td>
<td>2.353</td>
<td>-0.43</td>
<td>0.001</td>
<td>0.068</td>
<td>0.223</td>
<td>9</td>
</tr>
<tr>
<td>Sichuan</td>
<td>0.553</td>
<td>-0.183</td>
<td>0.649</td>
<td>-0.765</td>
<td>0.296</td>
<td>0.163</td>
<td>10</td>
</tr>
<tr>
<td>Zhejiang</td>
<td>1.247</td>
<td>-0.614</td>
<td>-0.154</td>
<td>0.121</td>
<td>0.026</td>
<td>0.161</td>
<td>11</td>
</tr>
<tr>
<td>Heilongjiang</td>
<td>-0.494</td>
<td>-0.289</td>
<td>0.739</td>
<td>-0.814</td>
<td>1.919</td>
<td>0.132</td>
<td>12</td>
</tr>
<tr>
<td>Hunan</td>
<td>0.457</td>
<td>-0.482</td>
<td>0.613</td>
<td>-0.698</td>
<td>0.058</td>
<td>0.036</td>
<td>13</td>
</tr>
<tr>
<td>Hubei</td>
<td>0.477</td>
<td>-0.22</td>
<td>0.136</td>
<td>0.02</td>
<td>-0.444</td>
<td>0.034</td>
<td>14</td>
</tr>
<tr>
<td>Xinjiang</td>
<td>-1.21</td>
<td>0.932</td>
<td>0.338</td>
<td>0.957</td>
<td>-0.926</td>
<td>-0.011</td>
<td>15</td>
</tr>
<tr>
<td>Anhui</td>
<td>0.182</td>
<td>-0.037</td>
<td>0.299</td>
<td>-0.242</td>
<td>-0.799</td>
<td>-0.057</td>
<td>16</td>
</tr>
<tr>
<td>Jilin</td>
<td>-0.538</td>
<td>-0.787</td>
<td>0.453</td>
<td>-0.42</td>
<td>1.452</td>
<td>-0.067</td>
<td>17</td>
</tr>
<tr>
<td>Guangxi</td>
<td>-0.314</td>
<td>-0.498</td>
<td>0.875</td>
<td>0.08</td>
<td>-0.705</td>
<td>-0.103</td>
<td>18</td>
</tr>
<tr>
<td>Shanghai</td>
<td>0.686</td>
<td>-1.293</td>
<td>-1.08</td>
<td>0.052</td>
<td>1.56</td>
<td>-0.113</td>
<td>19</td>
</tr>
<tr>
<td>Jiangxi</td>
<td>-0.113</td>
<td>-0.287</td>
<td>0.492</td>
<td>-0.343</td>
<td>-0.548</td>
<td>-0.126</td>
<td>20</td>
</tr>
<tr>
<td>Shaanxi</td>
<td>-0.204</td>
<td>0.001</td>
<td>-0.181</td>
<td>-0.383</td>
<td>0.016</td>
<td>-0.145</td>
<td>21</td>
</tr>
<tr>
<td>Yunnan</td>
<td>-0.391</td>
<td>0.01</td>
<td>0.432</td>
<td>-0.621</td>
<td>-0.531</td>
<td>-0.179</td>
<td>22</td>
</tr>
<tr>
<td>Fujian</td>
<td>0.528</td>
<td>-0.556</td>
<td>-0.382</td>
<td>-0.593</td>
<td>-0.29</td>
<td>-0.211</td>
<td>23</td>
</tr>
<tr>
<td>Chongqing</td>
<td>-0.166</td>
<td>-0.996</td>
<td>0.293</td>
<td>-0.263</td>
<td>0.202</td>
<td>-0.22</td>
<td>24</td>
</tr>
</tbody>
</table>
The economic development scores of each province and city are calculated and ranked by the regional economic composite factor score function. The results show that: in component $E_1$, Guangdong, Jiangsu and Shandong provinces rank relatively high, indicating that these regions have a high contribution to China’s economic development and help to improve China’s economic development; in component $E_2$, Shanxi, Hebei and Inner Mongolia provinces have the largest industrial contribution value, indicating that these regions mainly focus on industrial development and are typical industrial-driven provinces; in component $E_3$, Tibet, Tianjin, and Shanghai have the largest energy consumption share, indicating that these provinces and cities mainly focus on energy as the driving force of economic development and are typical energy-driven provinces and cities; in component $E_4$, Beijing has a larger share of the score than other provinces and cities, which reflects that Beijing has a strong foreign investment attractiveness and its private investment has become the main force driving China’s investment growth; in component $E_5$, the provinces such as Heilongjiang and Liaoning have higher scores for the social factor, indicating that these provinces have a positive employment trend, strong livelihood protection, and high overall consumption levels of residents.

4.5 Analysis of Spatial and Temporal Differences and the Chinese Experience

4.5.1 Statistical Analysis of Spatial and Temporal Differences

Considering the existence of negative values of the composite score index, we add 1 to the corresponding composite score values of each province and city to make all values positive, and also plot the composite score visualization of 31 provinces and cities in mainland China to better analyze the spatial differences of China’s regional economic growth during 2013-2019.

Combined with the calculation results of the composite factor score function, it can be seen that the top five provinces in the composite score ranking of the principal component factors are Guangdong, Shandong, Jiangsu, Liaoning, and Hebei, indicating that these five provinces have a high level of comprehensive economic development; while Hainan, Qinghai, and Tibet rank relatively low, indicating a low level of comprehensive economic development. Looking at the visualization of the overall composite score index distribution of Chinese provinces and cities in Figure 5, it can be found that the regional economic growth level of China generally shows a “three-step” decreasing trend from east to west during the period of 2013-2019. The eastern coastal region, mainly Guangdong, Shandong and Jiangsu, occupies most of the space for China’s export-oriented economic development due to its superior geographical location, and has broad market prospects for technology, capital and foreign investment. In addition, most of the eastern regions of China are also important hubs connecting China’s inland industries to the markets of foreign companies. Especially during the transition period from the 12th Five-Year Plan to the 13th Five-Year Plan, the government’s policy support for the eastern coastal areas
has increased, which to a certain extent has contributed to the high-quality development of China’s coastal economic zone. The central region occupies the majority of the Yangtze River Economic Belt in China, which enables economic and technological exchanges with the outside world to a certain extent, and many provinces and cities in the central region are the gathering places of China’s high-tech industries, which have strong industrial innovation and synergy capabilities, thus contributing more to the overall economic development than the western region. Due to its remote location, the western region of China largely restricts the external exchange and development of the regional economy, which also leads to the overall economic competitiveness of the western region being lower than that of the central and eastern regions.

As shown in Figure 6, the time trend of the composite score of Chinese provinces and cities shows that the overall level of China’s regional economic growth shows an increasing and then decreasing trend during 2013-2019, with an increasing phase during 2013-2014 and a decreasing phase during 2015-2019, which indicates that these seven years the imbalance of China’s regional economic development gradually increased. From the visualization of the composite scores of provinces and cities in Figure 7, it can be seen that the differences in economic development between eastern, northeastern and central parts of China changed more obviously between 2013 and 2019, and the corresponding composite score values all decreased, which is mainly related to resource allocation, industrial structure, regional openness and human capital factors.
Figure 6: Time trend of composite score by provinces and cities in China, 2013-2019
4.5.2 Chinese Experience and Empirical Evidence

- In terms of resource allocation, the strengths and weaknesses of inter-regional resource allocation can influence the speed and quality of economic development. However, the unbalanced distribution of resources prevails in most regions of China, which to some extent plays a decisive role in the type and distribution of industries in China, which in turn affects the overall economic development level.

- Considering the industrial structure, with the change of industrial types and cluster locations among Chinese regions, it has influenced the economic development dynamics of most of the eastern, northeastern and central regions, which in turn has accelerated the shift of economic and industrial centers of gravity. The industrial bases, mainly in the Beijing-Tianjin-Hebei Industrial Zone, Liaoning-Zhongnan Industrial Zone and Shanghai-Nanjing-Hangzhou Industrial Zone, have developed toward energy-intensive and high-energy-consuming industries, while the Pearl River Delta Economic Zone has developed toward technology-intensive and outward-oriented industries, mainly in high-tech information technology, which has caused the differences in industrial patterns and clusters between the north and the south of China.

- From the perspective of regional development, China has started to move from the top-level strategic concept to the practical cooperation stage in the 13th Five-Year Plan period, but
due to the international environment and industrial structure changes, China is facing downward pressure on the economy as a whole, and the market environment has also undergone formal changes, making China’s regional economic growth resistant.

In the context of human capital, on the one hand, during the period 2013-2019 China’s economic development achieved a shift from the deceleration and gear shift phase with rising capital contribution and declining total factor productivity to the new development concept phase; on the other hand, during this period, the number of labor force population in eastern and some central regions of China showed a declining trend, and the total labor input factor year by year. On the other hand, the number of the labor force in eastern and some central regions of China showed a decreasing trend during this period, and the total labor input factors decreased year by year, which in turn reduced the labor transfer effect and made the contribution of total factor productivity to economic growth brought by labor transfer decrease.

5 CONCLUSION

This paper measures the economic growth levels of 31 provinces and cities in inland China and analyzes their influencing factors. Combined with the results of factor analysis, we can see that: (1) there are obvious differences in China’s regional economic development. The overall economic development level of the eastern coastal region is higher than that of the western inland region, and the overall trend is a “decreasing ladder” among the eastern, central and western regions, and there is a certain gap between the regions. (2) The economic contribution rate of 31 provinces and cities was measured by the factor composite score function, and the results show that Guangdong, Shandong and Jiangsu have a higher level of economic development and rank in the top three in China, while Hainan, Qinghai and Tibet rank in the bottom three. (3) The characteristic roots, variance contribution rates and cumulative contribution rates of the 14 indicators selected in the article were measured by using principal component analysis, and the results showed that the level of economic development and social living standards are the primary factors affecting the degree of regional economic development. (4) The statistical analysis of spatial and temporal differences shows that four factors, namely, resource allocation, industrial structure, regional openness and human capital, are the key to the imbalance of China’s regional economic growth and the breakthrough point for China to achieve high-quality economic development in the future.

In response to the above findings, this paper makes the following recommendations.

First, optimize the industrial structure of each region, and promote the rationalization, coordination and ecological development of economic industries among regions. For regions with higher degree of economic development, strengthen the support at the level of innovation and technology, and promote industrial transformation and upgrading with technological breakthroughs; for regions with lower degree of economic development, formulate reasonable development plans according to the regional resource and environmental bearing situation, develop low-carbon and green development programs, and realize the green development of industrial economy.

Second, in view of the differences in economic development between different regions, actively strengthen the economic cooperation and connection between regions. To strengthen
the synergy of economic structure, the advantageous regions should focus on maintaining their development advantages and constantly updating innovative technologies to break through the development bottlenecks; the economically backward regions should take the initiative to change and break the traditional economic model, actively learn from the development experience of the excellent regions, rationalize the development measures, and maximize the leading effect of policies to promote the synergistic development of inter-regional economy in order to narrow the development gap in China as much as possible.

Thirdly, during the 14th Five-Year Plan period, we should pay attention to inter-regional science and technology linkages and improve the level of human capital to empower China’s high-quality economic development. Secondly, we should innovate a high-quality economic structure and build a modern market economy system to create an excellent environment and opportunities for China’s high-quality and coordinated economic development and to strengthen the foundation of China’s regional economic synergy. The last thing is to improve economic infrastructure construction, optimize the layout of industrial structure, curb the coexistence of high development and high energy consumption, and lay a solid foundation for the realization of high-quality economic development.

REFERENCES


The Construction of Internal Tax Audit System of Private Enterprises Based on Python

He Li
lihe6616@163.com
School of Public Finance and Administration, Harbin University of Commerce, Harbin City, Heilongjiang Province, China

Abstract: With the background of digital economy era, China's private enterprises have entered a critical period of transformation and upgrading. It is necessary to enhance the internal management and development of enterprises with the help of digitalization and informatization, avoid and deal with many operational risks, and promote the high-quality development of private economy. Therefore, this paper takes the internal tax audit management of private enterprises as the research object, makes full use of the functions and characteristics of network information technology, data mining technology and computer application technology, and relies on data processing class libraries such as Numpy and Pandas in Python environment to build a Web-based internal tax audit system of private enterprises. The whole system adopts B/S architecture design, and the Web Server is built by Flask framework, and the construction of various business logics and the deployment of API data interfaces are completed according to MVC pattern. The function setting of the system will comprehensively cover the whole process of internal tax audit work of private enterprises, which not only meets the needs of current enterprises and auditors, but also can give full play to the working efficiency of data algorithm models such as K-means and Apriori under data mining technology. The construction of the system not only greatly improves the efficiency of internal tax audit of private enterprises, but also promotes the informatization and intelligent development of internal management of enterprises.

Keywords: Private Enterprises, Internal Tax Audit, Python, Data Mining, Computer Mining.

1 INTRODUCTION

As an important part of China's national economy, private enterprises have played an irreplaceable role in economic development, improving people's livelihood, promoting innovation, deepening reform and opening up, and they are the backbone of China's socialist modernization. Since the 19th National Congress of the Communist Party of China, China's private economy has stood at a new historical starting point. It should actively grasp the theme of the times of high-quality development, adhere to the digital economy as the key driving force for transformation and upgrading, and empower the internal operation and management of enterprises through digitalization and informatization upgrading, which can effectively improve the backward problem of enterprise operation risk management mode, improve management efficiency, and ensure the smooth realization of organizational goals. [5]
The internal tax audit is an important part of many audit activities of enterprises, and it is also a necessary means for enterprises to take the initiative to prevent corporate tax risks. The internal tax audit inherits the basic characteristics of the internal audit work of enterprises, and is completely independent of other routine economic activities of enterprises such as accounting and finance. It aims at accounting and auditing the amount of tax paid by enterprises and the amount of tax payable on various incomes through standard audit methods, so as to ensure that enterprises pay taxes in accordance with the tax law in the process of tax payment, and make reasonable tax reduction or exemption. As the front node of the external tax audit, the internal tax audit of an enterprise's level of prevention and control is not only related to the immediate economic benefits of the enterprise, but also can provide the necessary data support for the future development of the enterprise, thus helping the enterprise to make timely decisions and adjustments within the scope of tax policies, and promote the rapid improvement of its own market competitiveness. However, the internal tax audit of private enterprises still faces many difficulties. Firstly, due to the particularity of private enterprises and their own limitations, there are obvious deficiencies in the establishment of management system and the handling of personnel hierarchy; Secondly, the lack of corresponding policies and regulations to support and guide, so that the internal tax audit work is ignored; Thirdly, the professional quality of internal tax auditors in private enterprises is not high, the audit mode is backward, and the routine on-site audit efficiency is low, which makes it more difficult to cope with the current complex electronic data and information, resulting in high labor cost. In view of this, this paper holds that private enterprises should adhere to innovation drive, actively implement the best operation mode of internal tax audit, and make full use of the application advantages of network information technology, data mining technology and computer application technology to build a Web-based internal tax audit system for private enterprises, and put forward a set of feasible solutions for internal tax audit of private enterprises. This system will cover all aspects of audit demand analysis, audit plan formulation, audit plan execution, audit report processing, etc., and apply K-means, Apriori and other data analysis and processing algorithm models under data mining to it, which will greatly improve the working efficiency of internal tax audit, speed up the information construction process and promote the scientific development of internal tax audit in private enterprises.

2 INTRODUCTION OF KEY TECHNOLOGIES

2.1 Web Technology

The essence of Web is a service program running on the Internet, and it is a distributed system. The overall operation of the Web architecture depends on the client and the server, that is, the user makes a "request" to the server through the client and parses the received content; The server is used to save content and "respond" to requests from browsers. According to the Web architecture, all technologies applied in the construction process are collectively referred to as Web technologies.

2.2 Python

The Python is a high-level scripting language that combines interpretive, compiler, interactive and object-oriented. As for Web application development, Python has many mature Web
development template technologies. As a lightweight Web framework, Flask framework is more flexible and extensible than other frameworks. The operation process of Flask framework is shown in Figure 1. When a user sends an HTTP request to the Server through the client browser, WSGI Server forwards the request to WSGI App, which completes logical processing of the request and returns the processing result to the server. The WSGI App can include multiple stacked middleware, which need to implement both the Server and the App at the same time, so it can play a regulatory role between the WSGI server and the WSGI application: for the server, the middleware acts as the application, and for the application, the middleware acts as the server.

Figure 1: Flask operation flow chart (original)

2.3 Data Mining Technology

As a kind of computer science and technology, data mining is a processing method for big data, which aims to extract information and knowledge that people don't know in advance but have potential usefulness from a large number of, incomplete, noisy, fuzzy and random actual data. The whole process involves three important links: data preparation, data mining implementation, result expression and interpretation, as shown in Figure 2. Among them, the object of data mining, that is, the data source, needs to take into account the characteristics of large quantity, multi-type, authenticity and applicability, and its quality directly affects the accuracy of the final result. The main steps of data mining include: data acquisition, data preprocessing, model establishment and operation, and final result expression.

Figure 2: Basic flow chart of data mining (original)
The construction of data mining model is the core of the whole data mining work. The construction of data mining model corresponds to data analysis method, which not only determines the application direction of data mining results, but also determines the construction of data mining model. The common data analysis methods include classification analysis, prediction analysis, cluster analysis, valuation analysis and correlation analysis, as shown in Table 1.

### Table 1: Types of Common Data Mining Algorithms

<table>
<thead>
<tr>
<th>Classify</th>
<th>Data analysis method</th>
<th>Data mining model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guided data mining</td>
<td>Classification analysis</td>
<td>Decision Tree, Random Forest, Neural Network</td>
</tr>
<tr>
<td>Predictive parsing</td>
<td>Regression tree and rough set method</td>
<td></td>
</tr>
<tr>
<td>Valuation analysis</td>
<td>SVM, Bayesian method</td>
<td></td>
</tr>
<tr>
<td>Unsupervised data mining</td>
<td>Correlation analysis</td>
<td>Pearson correlation coefficient, Aprior algorithm</td>
</tr>
<tr>
<td>Cluster analysis</td>
<td>K-means clustering algorithm, pedigree clustering</td>
<td></td>
</tr>
</tbody>
</table>

#### 2.4 Development Process

According to the system development requirements and the use requirements of the above key technologies, complete the configuration and deployment of the development environment. The overall design and development environment of private enterprise internal tax audit system is divided into two parts. The first is to use data processing class libraries such as Numpy and Pandas to complete the construction and training of algorithm models related to data mining in Python language environment; The second is to use Flask framework to build a Web Server, and integrate and package each functional module to form a Web application.

Firstly, according to the business process and organizational structure of private enterprises, the tax-related departments and business links are determined, and the selection of tax internal audit data sources and the collection of basic data are completed. The data source is mainly composed of financial information (vouchers, accounts, tables), corporate tax information, some corporate basic information and internal management information. All kinds of target data are summarized through data copying and transmission, and stored in a unified Mysql database system. The data in the whole is in a state of disorder and disorder, which needs to be preprocessed to improve the data quality and provide convenience for subsequent data mining. The data preprocessing mainly includes data cleaning, denoising and basic operation, and finally the basic financial and tax details with relatively uniform format and type, as well as key indicators such as tax rate, quick ratio, asset-liability ratio and expense ratio are obtained.

Secondly, according to the data information, the development of tax risk data mining function module is completed. As the core function module of the system, on the one hand, it can analyze and mine massive data information, and quickly obtain the identification and investigation of tax risks; On the other hand, it can integrate matplotlib data function class library to complete the visual expression of the final result. During the actual development process, the underlying
operating system is Linux CentOS 4.7, and the compilation environment is Visual Studio 2019. The development environment is Python 4.2.0, and the Anaconda integrated distribution can be selected for installation and deployment to shorten the time. In addition, the deployment and import of NumPy and Pandas need the help of PyCharm tool. The construction of all kinds of algorithm models can be completely realized by Python code, as shown in Figure 3, which is the key code of K-means clustering algorithm.

![Implementation code of K-means clustering algorithm](original)

Finally, for the development of the whole system on the Web Server side, the web server chooses Nginx server, version Nginx/Windows -1.12.2, project development language Python 4.2.0, development tool PyCharm 2018.3.1 x64 and database MySQL 5.7 to complete the construction and support of the system database system. The whole server is implemented by Flask 2.0.1 framework. Through the introduction of the above key technical theories, the overall environment of the system development, the configuration of related software and tools are determined, and the technical feasibility of the overall project of the internal tax audit system of private enterprises is also clarified.

### 3 FUNCTION REALIZATION

#### 3.1 Initiation of Tax Audit

In this function module, the internal audit users of private enterprises can independently complete the comprehensive query of enterprise tax data. Compared with the decentralization of all kinds of financial statements, personnel information, production information and management information under the traditional audit procedures, the system can realize the centralized presentation of all kinds of data information. When the user initiates the tax audit, the system will automatically generate the work project process guide, which will be completed
online from project determination, planning, personnel determination and final audit notice issuance, thus reducing the complicated document exchange of the traditional audit mode and greatly improving the working efficiency of the internal tax audit.

3.2 Tax Risk Identification

In this function module, internal audit users can identify and check the tax risks of related projects, data information and departments according to the tax audit plan. For example, after the user selects the basic information of tax audit department, project, time, etc. in the system in turn, according to the data rules, the tax rate of the project, quick ratio, net interest rate of assets, expense ratio, gross profit margin and other indicators are taken as the key features, and the K-means clustering algorithm model is selected to complete the data classification, thus generating the classification standard, which can be compared according to the standard. If the two groups of data are similar, it can be judged that the tax risk is low, whereas the tax risk is high. Table 2 shows the algorithm formulas of some indexes, and Table 3 shows the variance analysis diagram of some K-means clustering algorithms, in which the significant difference value (Sig) is the criterion for judging whether the enterprise has tax risks, which can reduce the scope of the follow-up on-site audit and improve the efficiency.

<table>
<thead>
<tr>
<th>Number</th>
<th>Index name</th>
<th>Algorithm formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Tax rate</td>
<td>[ Q_1 = \frac{M_1(Tax \ increment)}{K(Actual \ sales \ revenue)} ]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[ Q_2 = \frac{M_2(Income \ tax \ amount)}{K(Actual \ sales \ revenue)} ]</td>
</tr>
<tr>
<td>2</td>
<td>Quick ratio</td>
<td>[ W = \frac{SS(quick \ assets)}{SF(current \ liabilities)} \times 100% ]</td>
</tr>
</tbody>
</table>

Table 2: Calculation formulas of some indexes

<table>
<thead>
<tr>
<th>Index</th>
<th>Cluster</th>
<th>Error</th>
<th>Test value (F)</th>
<th>Significant difference (Sig)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tax burden rate</td>
<td>0.000</td>
<td>1.0</td>
<td>0.091</td>
<td>0.761</td>
</tr>
<tr>
<td>Quick ratio</td>
<td>0.000</td>
<td>1.0</td>
<td>1.667</td>
<td>0.201</td>
</tr>
<tr>
<td>Asset load rate</td>
<td>19.414</td>
<td>1.0</td>
<td>120.611</td>
<td>0.000</td>
</tr>
<tr>
<td>Net interest rate of assets</td>
<td>18.477</td>
<td>1.0</td>
<td>58.564</td>
<td>0.000</td>
</tr>
</tbody>
</table>
3.3 Tax Audit Report

According to the tax risk identification and investigation results, auditors can conduct targeted on-site tax audits and make detailed records of tax problems found during the period. After the on-site audit is completed, the user will summarize the data according to the tax audit results and tax categories, and complete the compilation of the Tax Audit Statement in the system to form a preliminary audit opinion. The final tax audit report will be released in the system after the audit is correct.

4 CONCLUSIONS

This paper aims at promoting the internal tax audit mode of private enterprises, and makes full use of the functional characteristics of network information technology, data mining technology and computer application technology to build a Web-based internal tax audit system for private enterprises. The functions of the system will cover all aspects of audit demand analysis, audit plan formulation, audit plan execution, audit report processing, etc., with concise, convenient and efficient operation, the efficiency of internal tax audit will be greatly improved, the accuracy of tax risk investigation and judgment will be increased, the informatization construction of tax audit will be accelerated, and the scientific development of internal management of private enterprises will be promoted.

REFERENCES

The Moderating Effect of Digitalization on Diversification and Performance of Enterprises

Qiong Chen¹, *, Zhengmao Yang²

*Qiong Chen. Email: 760792544@qq.com¹

School of Business Administration, Zhejiang Gongshang University, Hangzhou, 310000, China¹
School of Computer Science and Technology, Zhejiang University, Hangzhou, 310000, China²

Abstract: Digital transformation has become an important strategic means for many enterprises to solve the problem of redundant resources and improve efficiency. In order to explore whether digital transformation can improve the performance of diversified enterprises, this paper collected a total of 6238 valid data from Chinese A-share listed companies from 2010 to 2020, made an empirical analysis with the level of enterprise digitalization as the adjusting variable, and drew the following conclusions: the degree of enterprise diversification is negatively related to enterprise performance, and the enterprise's hasty diversification will damage enterprise performance; the level of enterprise digitalization plays a negative role in regulating the relationship between enterprise diversification and enterprise performance. The improvement of enterprise digitalization can significantly curb the corresponding drawbacks of enterprise diversification, which is beneficial to the improvement of enterprise performance.

Keywords: Diversified Operation, Digital Transformation, Enterprise Performance.

1 INTRODUCTION

In recent years, with the development of artificial intelligence technology, blockchain technology, cloud computing technology, big data technology and other digital technologies, the original production and operation mode of enterprises has changed, and more and more enterprises have been attracted and promoted to enter the path of digital transformation. According to the White Paper on the Development of China's Digital Economy (2022) released by the China Academy of Information and Communication Research, since 2012, China's digital economy has maintained a growth rate of 15.9%. In 2021, the scale of China's digital economy will reach 45.5 trillion yuan, with a nominal year-on-year growth of 16.2%. Digitalization has become a major boost to China's economic development.

For enterprises, diversification can effectively reduce business risks and improve business performance. However, in recent years, more and more enterprises have proved from practice that diversified operation may not bring ideal performance, on the contrary, enterprises may fall into operational difficulties. This difference in the actual situation has aroused the attention of scholars. According to the principal-agent theory, when an enterprise adopts a diversification strategy, the number of managers needed to be managed will gradually increase, which increases the difficulty of the company's management and operation. At the
same time, each manager may make opportunistic mistakes out of their own development plans, which will damage the overall interests of the enterprise, and ultimately the actual performance of the enterprise will decline. Therefore, it is extremely urgent to solve the problems brought about by the diversification of enterprises.

So, can enterprises improve their digital level and carry out digital transformation effectively solve the problems brought about by diversified operation, and then increase their efficiency? With such a problem, based on the exploration of the relationship between enterprise diversification and enterprise performance, this paper considers whether the variable of enterprise digitalization plays a good role in regulating this relationship, whether it can strengthen the positive effect of diversification on enterprise performance, and better promote enterprise development.

2 LITERATURE REVIEW AND HYPOTHESIS PRESENTATION

2.1 Diversification and enterprise performance

Diversification Strategy refers to a development Strategy in which an enterprise operates two or more products or services with different basic economic purposes at the same time [6]. Since Rumelt raised the question of "specialization" or "Diversification" for enterprise operation in 1974, the research on this aspect has become a key topic for scholars to discuss. However, no matter at home or abroad, many scholars have not reached a unified conclusion on the merits of corporate diversification [7]. Whether China's listed companies diversify and the degree of diversification are affected by various internal and external factors [8]. Such as internal and external corporate governance mechanism [9], government intervention, etc. [10]. Of course, different scholars have different views on enterprise diversification, but generally it can be divided into three categories: diversification has neutral characteristics, diversification is conducive to enterprise performance, and diversification reduces enterprise performance.

Diversification Strategy refers to a development Strategy in which an enterprise operates two or more products or services with different basic economic purposes at the same time [6]. Since Rumelt raised the question of "specialization" or "Diversification" for enterprise operation in 1974, the research on this aspect has become a key topic for scholars to discuss. However, no matter at home or abroad, many scholars have not reached a unified conclusion on the merits of corporate diversification [7]. Whether China's listed companies diversify and the degree of diversification are affected by various internal and external factors [8]. Such as internal and external corporate governance mechanism [9], government intervention, etc. [10]. Of course, different scholars have different views on enterprise diversification, but generally it can be divided into three categories: diversification has neutral characteristics, diversification is conducive to enterprise performance, and diversification reduces enterprise performance.

Diversification can affect the profitability of enterprises [11]. Diversification has a positive effect on the value of enterprises. Diversification can improve the value of enterprises [8]. At the same time, diversification can also reduce the risk of enterprise operation [21], optimize the level of cash holdings [13], ease resource constraints [14], and ultimately improve enterprise income.
Compared with the view that diversification can bring more benefits and hold a moderate position on diversification, this paper believes that blind diversification is detrimental to the development of enterprises. Diversification decision-making behavior and enterprise performance are negatively correlated, which will lead to the discount of company value [9]. Although the traditional view is that enterprise diversification can disperse the risk of enterprise operation, Wu Guoding and Zhang Huili (2015) have shown through empirical research that diversification can not effectively disperse the financial risk of enterprises, and pointed out that many European and American enterprises have closed down under the background of the financial crisis, which is inseparable from diversification expansion. Wang Zhiqiang and Ren Zhenchao (2021) also confirmed through research that enterprise diversification will indeed reduce the solvency of enterprises. Anish et al.

At present, entrepreneurs have gradually become more rational about the possible "trap" of diversification [22], and returning to the main business has gradually become a new business trend [23]. Therefore, this paper proposes the following assumptions:

Hypothesis 1: There is a negative relationship between diversification level and enterprise performance.

2.2 Digital Transformation and Enterprise Performance

Enterprise digitalization is a process of organizational change [27]. It means that enterprises make comprehensive use of digital technology to continuously adjust the structure and operation of the organization in terms of value chain, business process and product and service innovation, so as to promote enterprises to increase income, improve business, replace or transform business processes, and create a digital business environment with digital information as the core [24]. With the development of informatization and digitalization, more and more enterprises have got rid of the original business model. Digital office, digital marketing and digital production and operation have gradually become the status quo of modern enterprises' daily operation.

Xie Lijuan (2019) supported the digital transformation of the retail industry and believed that the digital transformation could significantly improve the circulation efficiency of enterprises and improve their performance. Liu Donghui et al. (2022) believed that enterprises can accurately predict customer demand through big data and other digital technologies, effectively solve the supply and demand problem, and then improve enterprise performance. Qi Yudong and Cai Chengwei (2020) believe that digital transformation can solve the "long arm jurisdiction" effect caused by the development and growth of enterprises through online opening, sharing, connection, collaboration and other ways, so as to manage and operate enterprises more efficiently and improve the performance of enterprises.

It is a general trend for enterprises to carry out digital transformation, and it can effectively improve the operation and management capabilities of enterprises in many aspects, so as to improve enterprise performance and industry competitiveness. Therefore, the following assumptions are proposed in this paper:

Hypothesis 2: The digital level plays a negative role in regulating the relationship between corporate diversification and corporate performance.
3 METHODS

3.1 Sample selection and data collection

The data of listed companies in this paper mainly comes from the CSMAR (Guotai'an) database, and is sorted out in the following ways: (1) First, download the data of listed companies from the CNRDS database from 2010 to 2020; (2) Eliminate the financial and insurance industry; (3) Remove the incomplete samples of related variables. (4) Excluding the enterprises that were PT, ST, * ST and delisted during the study period, the data of these enterprises may be abnormal. After relevant sorting, 6238 valid sample observations were obtained.

3.2 Variable definition and measurement

(1) Interpreted variable

Corporate Performance (ROA). This paper refers to the common practice of previous scholars \[1-3\], and selects the return on total assets (ROA) in the enterprise financial indicators to measure (net profit/average balance of total assets).

(2) Explanatory variable

Degree of diversification (EI). This paper uses entropy index \[18\] to measure the degree of enterprise diversification. The calculation method is: entropy index (EI)=\[\sum P_i * \ln (1/P_i)\]. \(P_i\) is the \(i\)th main business income/total business income of the enterprise.

(3) Regulating variable

Digitization level (Di). This paper draws on the current common practice of scholars such as Wu Fei (2021) and Wang Molin (2022) to collect the frequency of keywords such as artificial intelligence technology, blockchain technology, cloud computing technology, big data technology, and digital technology applications from the annual reports of listed companies as indicators of the enterprise's digital level (Di).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Key word</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digital level</td>
<td>artificial intelligence technology</td>
<td>Artificial intelligence, business intelligence, image understanding, investment decision support system, intelligent data analysis, intelligent robot, machine learning, deep learning, semantic search, biometrics, face recognition, voice recognition, identity verification, automatic driving, natural language processing</td>
</tr>
<tr>
<td></td>
<td>Blockchain technology</td>
<td>Blockchain, digital currency, distributed computing, differential privacy technology, smart financial contract</td>
</tr>
<tr>
<td></td>
<td>Cloud computing technology</td>
<td>Cloud computing, stream computing, graph computing, memory computing, multi-party security computing, brain like computing, green computing, cognitive computing, fusion architecture, 100 million level concurrency, EB level storage, the Internet of Things, and information physics systems</td>
</tr>
</tbody>
</table>
(4) Control variable

In addition to the main variables, this paper also selects enterprise size, asset liability ratio (Lev), total asset turnover ratio (ATO), cash flow ratio (Cf), and operating income growth rate (Gro) as control variables. At the same time, this paper also sets year and industry as dummy variables.

<table>
<thead>
<tr>
<th>Variable Type</th>
<th>Variable name</th>
<th>Variable symbol</th>
<th>Variable definition and measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent variable</td>
<td>Enterprise performance</td>
<td>ROA</td>
<td>Net profit/average balance of total assets</td>
</tr>
<tr>
<td>Independent variable</td>
<td>Diversification level</td>
<td>EI</td>
<td>((EI) = \sum Pi \times \ln (1/Pi), Pi is the ith main business income/total business income of the enterprise)</td>
</tr>
<tr>
<td>Regulating variable</td>
<td>Digital level</td>
<td>Di</td>
<td>Collect the total frequency of keywords such as artificial intelligence technology, blockchain technology, cloud computing technology, big data technology, and digital technology in the annual reports of listed companies</td>
</tr>
<tr>
<td>Control variable</td>
<td>Enterprise scale</td>
<td>Size</td>
<td>The natural logarithm of the total assets of the enterprise</td>
</tr>
<tr>
<td></td>
<td>Debt level</td>
<td>Lev</td>
<td>Total liabilities at the end of the year/total assets at the end of the year</td>
</tr>
<tr>
<td></td>
<td>Turnover rate of total assets</td>
<td>ATO</td>
<td>Operating income/total average assets</td>
</tr>
<tr>
<td></td>
<td>Cash flow ratio</td>
<td>Cf</td>
<td>Net cash flow from operating activities/total assets</td>
</tr>
<tr>
<td></td>
<td>Growth ability</td>
<td>Growth</td>
<td>Operating income of the current year/operating income of the previous year - 1</td>
</tr>
</tbody>
</table>
3.3 Hypothetical model

The relationship between enterprise diversification and enterprise performance has not been verified, so the model (1) is established as follows:

$$\text{ROAi,t} = \beta_0 + \beta_1 \text{EIi,t} + \beta_2 \text{Sizei,t} + \beta_3 \text{Levi,t} + \beta_4 \text{ATOi,t} + \beta_5 \text{Cfi,t} + \beta_6 \text{Groi,t} + \sum \text{Year} + \sum \text{Industry} + \varepsilon_{i,t} \quad (1)$$

The moderating effect of digital level on diversification and corporate performance has not been verified, so models (2) and (3) are established as follows:

$$\text{ROAi,t} = \beta_0 + \beta_1 \text{EIi,t} + \beta_2 \text{Di}_{i,t} + \beta_3 \text{Sizei,t} + \beta_4 \text{Levi,t} + \beta_5 \text{ATOi,t} + \beta_6 \text{Cfi,t} + \beta_7 \text{Groi,t} + \sum \text{Year} + \sum \text{Industry} + \varepsilon_{i,t} \quad (2)$$

$$\text{ROAi,t} = \beta_0 + \beta_1 \text{EIi,t} + \beta_2 \text{Di}_{i,t} + \beta_3 \text{EIi,t} \cdot \text{Di}_{i,t} + \beta_4 \text{Sizei,t} + \beta_5 \text{Levi,t} + \beta_6 \text{ATOi,t} + \beta_7 \text{Cfi,t} + \beta_8 \text{Groi,t} + \sum \text{Year} + \sum \text{Industry} + \varepsilon_{i,t} \quad (3)$$

4 RESULTS

4.1 Descriptive statistics and correlation analysis

This paper uses Stata software to conduct descriptive statistics and VIF analysis on related variables, and the statistical results are shown in Table 3. It can be seen from the table that the average digital level (Di) is 1.584, and the variance is 1.437. Generally speaking, the digital level of listed companies in China is uneven, with large differences, and the overall level is low, which is still in the early stage of transformation. In addition, the VIF value of each variable is less than 0.5, so there is no obvious correlation.

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>mean</th>
<th>sd</th>
<th>min</th>
<th>max</th>
<th>VIF</th>
<th>1/VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROA</td>
<td>6238</td>
<td>0.034</td>
<td>0.082</td>
<td>-0.415</td>
<td>0.244</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>EI</td>
<td>6238</td>
<td>0.374</td>
<td>0.441</td>
<td>0</td>
<td>2.117</td>
<td>1.05</td>
<td>0.954</td>
</tr>
<tr>
<td>Di</td>
<td>6238</td>
<td>1.584</td>
<td>1.437</td>
<td>0</td>
<td>5.753</td>
<td>1.02</td>
<td>0.981</td>
</tr>
<tr>
<td>Size</td>
<td>6238</td>
<td>22.22</td>
<td>1.194</td>
<td>19.716</td>
<td>26.395</td>
<td>1.37</td>
<td>0.728</td>
</tr>
<tr>
<td>Lev</td>
<td>6238</td>
<td>0.411</td>
<td>0.194</td>
<td>0</td>
<td>0.99</td>
<td>1.46</td>
<td>0.683</td>
</tr>
<tr>
<td>ATO</td>
<td>6238</td>
<td>0.627</td>
<td>0.387</td>
<td>0.044</td>
<td>2.777</td>
<td>1.11</td>
<td>0.897</td>
</tr>
<tr>
<td>Cf</td>
<td>6238</td>
<td>0.048</td>
<td>0.064</td>
<td>-0.224</td>
<td>0.258</td>
<td>1.1</td>
<td>0.905</td>
</tr>
<tr>
<td>Gro</td>
<td>6238</td>
<td>0.208</td>
<td>0.494</td>
<td>-0.732</td>
<td>4.806</td>
<td>1.03</td>
<td>0.968</td>
</tr>
</tbody>
</table>
4.2 Correlation analysis

Under the preliminary exploration of correlation, we can know that the coefficient between enterprise diversification (EI) and enterprise performance (ROA) is -0.114, and is significantly correlated at the level of 0.01. On the surface, with the continuous improvement of diversification, the performance of enterprises gradually declines, and diversification has a negative impact on enterprise performance, which is consistent with our hypothesis 1. The coefficient between digitalization level (Di) and enterprise performance is -0.032, and there is a negative correlation between digitalization and performance, which is contrary to our preliminary assumption, but the final result should refer to regression analysis.

<table>
<thead>
<tr>
<th></th>
<th>ROA</th>
<th>Int</th>
<th>Di</th>
<th>Size</th>
<th>Lev</th>
<th>ATO</th>
<th>Cf</th>
<th>Gro</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROA</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EI</td>
<td>0.114**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Di</td>
<td>-0.032**</td>
<td>0.062**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size</td>
<td>0.028**</td>
<td>0.151**</td>
<td>0.021*</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lev</td>
<td>0.318**</td>
<td>0.120**</td>
<td>0.037**</td>
<td>0.493*</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATO</td>
<td>0.194**</td>
<td></td>
<td>0.068**</td>
<td>0.126*</td>
<td>0.185*</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cf</td>
<td>0.350**</td>
<td>0.103**</td>
<td>0.041**</td>
<td>0.022*</td>
<td>0.194*</td>
<td>0.133*</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Gro</td>
<td>0.257**</td>
<td>0.001</td>
<td>0.049**</td>
<td>0.053*</td>
<td>0.001</td>
<td>0.159*</td>
<td>-0.002</td>
<td>1</td>
</tr>
</tbody>
</table>

*t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

4.3 Regression analysis

It can be seen from the regression results of model (1) that there is a negative correlation between enterprise diversification level (EI) and enterprise performance (ROA), and there is a significant correlation at the level of 0.01. On the surface, with the improvement of enterprise diversification, enterprise performance is damaged. Hypothesis 1 is tenable. According to model (3), the coefficient of EI * Di, the product of diversification and digitalization, is -0.003, and is significantly correlated at the level of 0.05, indicating that digitalization negatively regulates the relationship between diversification and enterprise performance. With the improvement of enterprise digitalization, the enterprise has made great progress in information communication, personnel management, marketing and other aspects, effectively solving the problems of miscellaneous multi-party management, poor communication and other problems caused by diversified operation, and ultimately reducing the phenomenon of
performance decline caused by the adoption of diversification strategy and even making the performance gradually rise. Hypothesis 2 is established.

Table 5 Regression Analysis

<table>
<thead>
<tr>
<th>variable</th>
<th>ROA (1)</th>
<th>ROA (2)</th>
<th>ROA (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EI</td>
<td>-0.011***</td>
<td>-0.011***</td>
<td>-0.006**</td>
</tr>
<tr>
<td></td>
<td>(-5.531)</td>
<td>(-5.378)</td>
<td>(-2.067)</td>
</tr>
<tr>
<td>Di</td>
<td>-0.003***</td>
<td>-0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-3.098)</td>
<td>(-0.977)</td>
<td></td>
</tr>
<tr>
<td>EI*Di</td>
<td>-0.003**</td>
<td></td>
<td>-0.003**</td>
</tr>
<tr>
<td></td>
<td>(-2.191)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size</td>
<td>0.014***</td>
<td>0.014***</td>
<td>0.014***</td>
</tr>
<tr>
<td></td>
<td>(16.281)</td>
<td>(16.352)</td>
<td>(16.275)</td>
</tr>
<tr>
<td>Lev</td>
<td>-0.162***</td>
<td>-0.163***</td>
<td>-0.164***</td>
</tr>
<tr>
<td></td>
<td>(-30.476)</td>
<td>(-30.654)</td>
<td>(-30.724)</td>
</tr>
<tr>
<td>ATO</td>
<td>0.036***</td>
<td>0.036***</td>
<td>0.036***</td>
</tr>
<tr>
<td></td>
<td>(15.290)</td>
<td>(15.538)</td>
<td>(15.294)</td>
</tr>
<tr>
<td>Cf</td>
<td>0.320***</td>
<td>0.317***</td>
<td>0.316***</td>
</tr>
<tr>
<td></td>
<td>(22.786)</td>
<td>(22.511)</td>
<td>(22.454)</td>
</tr>
<tr>
<td>Gro</td>
<td>0.034***</td>
<td>0.035***</td>
<td>0.035***</td>
</tr>
<tr>
<td></td>
<td>(19.481)</td>
<td>(19.634)</td>
<td>(19.668)</td>
</tr>
<tr>
<td>Constant</td>
<td>8.031***</td>
<td>7.414***</td>
<td>7.386***</td>
</tr>
<tr>
<td></td>
<td>(9.337)</td>
<td>(8.402)</td>
<td>(8.372)</td>
</tr>
<tr>
<td>Year</td>
<td>Control</td>
<td>Control</td>
<td>Control</td>
</tr>
<tr>
<td>Industry</td>
<td>Control</td>
<td>Control</td>
<td>Control</td>
</tr>
<tr>
<td>Observations</td>
<td>6,238</td>
<td>6,238</td>
<td>6,238</td>
</tr>
<tr>
<td>R²</td>
<td>0.321</td>
<td>0.322</td>
<td>0.323</td>
</tr>
<tr>
<td>Adj-R²</td>
<td>0.321</td>
<td>0.321</td>
<td>0.322</td>
</tr>
</tbody>
</table>

The t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

5 CONCLUSION

Digital transformation has become an important means for many enterprises to increase production and efficiency, especially for diversified enterprises. Through the collection of keywords and relevant data of digital transformation of listed companies, the empirical study confirms that diversification is harmful to enterprise performance, but the improvement of digital level is helpful to improve this situation and improve enterprise performance. This study has certain reference significance for enterprises in digital transformation and waiting.

REFERENCES


Studies on Financial Center Selection Using Spatial Statistical Analysis and Combination Evaluation

Longxiao Yuan¹, Yuxin Jia²,*

Email: 2482516799@qq.com¹, *Email: 2360919540@qq.com²

University College London, London, the United Kingdom¹
University of Oxford, Oxford, the United Kingdom²

Abstract: For a long time, due to the level of economic development and natural resource endowment, China has had an obvious spatial-structural non-equilibrium. As an essential part of the economic system, finance exhibits relatively solid regional characteristics. Financial resources and related factors driven by them rapidly gather in the central cities, resulting in the gradual formation of regional financial centers. Financial centers realize the optimal allocation of available resources and drive regional economic growth through financial agglomeration, expansion diffusion, spillover, and economic growth. In this paper, spatial statistical analysis is successfully applied to the issue of financial agglomeration in China. The research results indicate that although China's regional economic growth has a statistically significant spatial aggregation distribution, the spatial aggregation distribution for financial development demonstrates a relatively insignificant level. Economic agglomeration far precedes the financial counterpart. The spatial agglomeration level of China's regional financial development is persistently increasing, as more-developed areas have particular driving effects on the financial development of surrounding provinces and cities. This pattern reflects the huge supporting role of regional financial centers on the economy and finance of surrounding areas.

Keywords: Financial Center, Spatial Statistical Analysis, Combined Evaluation, Mann-Kendall Test.

1 INTRODUCTION

At present, there are many financial centers of different types, levels, and scales distributed all over the world. By observing these financial centers' development status or comparing their evolution process, there is limited theory of of financial centers due to their diverse historical backgrounds, qualifications, and endowments. However, regarding the economic impact, all financial centers have the same objective to exhibit radiation effects on the region or country with uneven agglomeration. At the same time, all regional financial centers have strong spatiotemporal characteristics, for they are products of their times and have their unique evolution process. Being the financial hub of the region, they play the role of the central node.
The research object of this paper is the regional financial centers in Mainland China. Their formation and development also have apparent temporal and spatial characteristics.

Mining its spatiotemporal evolution mechanism can promote the coordinated development of regional finance and is conducive to the optimal allocation of financial resources. At the same time, as an emerging market, China and its construction and development of financial centers are crucially influenced by the international environment and the domestic situation. Based on this, this research focuses on the background of the following two aspects.

On the one hand, China's regional financial centers present solid spatiotemporal characteristics. The regional financial center is a spatial concept and plays a pivotal role in gathering, distributing, and amalgamating financial resources within a relatively complete geographical unit. In a country with a vast territory and long-term unbalanced regional economic development, the spatial characteristics of regional financial centers are more prominent.

First, the formation and distribution of China's regional financial centers are spatially different. Due to the significant differences in natural endowments, economic development, and institutional factors across China, the formation mechanism of China's regional financial centers also has spatial differences. Some regional financial centers are naturally formed with their natural resource endowment as their advantage and gradually centered on their geographical location. Some are based on differences in economic development, formed gradually, with market agglomeration as the driving force. Others are based on the government's positioning as a financial center formed through government leadership. The spatial differences in the formation mechanism also make China's regional financial centers unbalanced: the distribution of regional financial centers in the eastern region is relatively dense, while those in the Midwest and Northeast are not dominant in quantity and scale.

Secondly, the spatial effects of China's regional financial centers are also different. Regional financial centers of different scales and levels have different agglomeration and radiation effect ranges. Its effect on allocating financial resources and upgrading industrial structure in the region has a strong spatial dependence. There are also asymmetric two-way spillover effects between regional financial centers and other spaces\[6\]. China's regional financial centers also have prominent time characteristics. Due to the staged differences in the economic development of various regions in China, the major regional financial centers also differ significantly in the timing of their formation. As early as the end of the 1920s, Shanghai was already a regional financial center with world influence. Meanwhile, Beijing's financial influence can also radiate to the whole country. At the same time, the construction and development of the ASEAN Free Trade Area accompany the rise of the Nanning regional financial center, nearly a century behind Beijing and Shanghai in terms of time. Many regional financial centers in China have gradually developed and grown, relying on different locations and economic advantages after the economic reform. Due to the difference in formation time, the stages of development of China's regional financial centers are also quite different. According to the latest Global Financial Center Index (GFCI) made by Z/Yen Group, Shanghai, Beijing, Shenzhen, Qingdao, and Dalian have been considered international financial centers, while Guangzhou and Tianjin serve as secondary financial centers. However, most regional financial centers in mainland China are still in the construction and development stage. Although they have clarified their development ideas and issued supportive policies, there is still a long way to go to realize the function of an actual financial center.
The future development of China's regional financial centers will also be very different. Currently, most of China's regional financial centers have gotten rid of blind imitation and disorderly competition and set up corresponding target positioning according to their qualifications and endowments. In the future, the development of regional financial centers will be more differentiated. Moreover, form a situation of misplaced development and healthy competition. However, it is not excluded that some financial centers will decline and eventually be eliminated in future competition, and new centers will take advantage of the trend.

On the other hand, the new international and domestic situation has a far-reaching impact on regional financial centers. The global financial turmoil triggered by the US subprime mortgage crisis has been nearly ten years old. However, the global economy and finance have yet to recover. The latest World Economic Outlook and the Global Financial Stability Report published by the International Monetary Fund (IMF) published in 2016 issued similar warnings about a weak economy, weak recovery, and rising risks to financial stability. The world economy has not yet emerged from the shadow of the financial crisis. We are still in the post-financial crisis era. This crisis has profoundly affected and changed the international financial landscape. As the global economic and financial center of gravity" move east," the status of financial centers in East Asia is rising. In particular, the financial center of mainland China has developed rapidly. The post-financial crisis also gave birth to the fourth industrial revolution, driven by the Internet of Things technology, big data and cloud computing, artificial intelligence, and 3D printing technology. The new revolution of production, service, and life informatization and intelligence. At the same time, also heralds the arrival of a new era of financial reform. In particular, big data and mobile Internet finance have quietly changed the operation and supervision system of modern finance' Facing such unprecedented challenges and opportunities [10]. The development of China's regional financial centers has a long way to go. Facing the sluggish recovery of the global economy, they should better play the function of financial resource circulation and integration. Facing the rise of East Asian financial centers, they should strengthen inter-regional financial cooperation, mutual benefit, and win-win. They must be determined to innovate, improve the efficiency of financial operations, and enhance financial regulatory functions.

In recent years, with the steady development of China's economy, China has become the third-largest economy in the world, and China's economic strength has supported multiple economic centers, which drive China's economy forward. At the same time, the construction of financial markets in many cities has gradually improved, and they have the conditions to build international financial centers or regional financial centers. Moreover, many cities have successively issued calls for the construction of financial centers at different levels through different means, and the construction of financial centers in various cities is proceeding in an orderly manner. All signs indicate that our country has the conditions to establish multiple financial centers. However, the construction of a financial center is a multi-level, multi-objective, multi-factor, and complex-structured systematic project [1]. Some things could be improved in building financial centers in multiple cities simultaneously. Competition and waste of resources Various financial center cities tend to adopt preferential policies similar to investment promotion competitions, overemphasize extensive preferential conditions, and ignore their characteristics. Local governments at all levels may pay too much attention to the number of financial institutions, Ignoring the creation of a peaceful financial environment [9].
In addition, China has a vast land area, objective differences exist in the economic development of different regions, and the imbalance in financial development is particularly prominent. The construction of a financial center also requires the state's macro-control objectively. In forming a financial center, the strategic orientation of the central government has an important impact on the healthy development of the city. This paper creatively applies the method of spatial statistical analysis and combined evaluation to the research of China's regional financial centers. The results may have good reference value for the national strategic choice in the process of China's financial center construction.

2 LITERATURE REVIEW

The concept of "Agglomeration" comes from "Industrial Agglomeration," first proposed by Alfred Marshall, referring to the process in which a particular industry is concentrated to a relatively high degree in a particular area, and the elements of industrial capital are continuously gathered. However, early studies on industrial agglomeration mainly focused on manufacturing and other industrial fields. As a particular financial service industry, the rise of the financial agglomeration phenomenon has naturally aroused extensive attention and research in academic circles. However, the academic circle has not yet reached a unified definition regarding the concept and connotation of "financial agglomeration". Moreover, although financial agglomeration belongs to the category of industrial agglomeration, the financial service industry, as a particular industry, has unique attributes (including high mobility, economic core, and dominance) that make its agglomeration connotation different from general. There are essential differences in industrial agglomeration, mainly manifested in the degree of agglomeration, content, speed, motivation, and scope of influence (Huang Jieyu & Yang Zaibin, 2006). Most theoretical research abroad is about analyzing the connotation of industrial agglomeration. Marshall (1890) put forward the concept of industrial agglomeration for the first time in his classic book "Principles of Economics," and based on external economies of scale. The theoretical thought of the concept defines its connotation as being in a specific area.

Reasonably measuring the level of financial agglomeration is the primary content of studying financial agglomeration. Scholars have fully used appropriate research methods on industrial agglomeration and used different indicators to measure the degree of financial agglomeration, which better reflects the spatial layout characteristics of the financial industry within a specific range. Ding Yi et al. (2009) first combined the influencing factors of the financial industry to screen and establish an evaluation index system for the level of financial agglomeration. Based on the economic data at the provincial level in China, they used principal component analysis to measure the degree of financial agglomeration, which reflected more comprehensively. The degree of financial agglomeration in various provinces and municipalities in China. It not only measures the degree of financial agglomeration in each region but also analyzes the mutual influence of each region on the degree of financial agglomeration. Ren Yinghua (2010) took the financial service industry in Lujiazui, Shanghai, as an example and built an evaluation index system for financial agglomeration core capabilities. Combining the AHP and the characteristics of financial formats, he designed a fuzzy comprehensive evaluation model to measure the core capabilities of financial agglomeration.
3 METHODS AND RESULTS

Spatial statistics emerged in the 1970s, and its basic idea is to introduce the interdependence of individuals, such as regions or institutions, into the model. For more than 30 years, foreign climatologists, geographers, economists, scientists, and anthropologists have been widely involved in the research and application of this method [7]. In the 1990s, Zhang described spatial statistics. However, the popularization and application of spatial statistics in China are relatively slow due to the constraints of calculation and other aspects. Until recent years, the breakthrough and popularization of foreign spatial statistical software calculations allow more and more domestic scholars to begin to join in the study, research, and application of spatial statistics.

Spatial statistical analysis refers to the statistical analysis of spatial data, which establishes the statistical relationship between data through spatial location, and understands data related to geographical location, spatial dependence, spatial autocorrelation, or spatial association. The spatial statistical analysis method believes that a specific geographical phenomenon or a particular attribute value on a regional unit is usually related to the same phenomenon or attribute value on an adjacent regional unit [3].

Almost all spatial data have spatial dependence or spatial autocorrelation characteristics. The spatial autocorrelation coefficient is the primary index to measure spatial autocorrelation. Commonly used methods for testing global spatial autocorrelation include the global Moran I test (global Moran I index), the Geary test (Global Geary C coefficient), and the Getis test (global Getis index). Methods for testing local spatial autocorrelation include the local Moran I test (local Moran I index), local Geary test, local Getis test, and Moran scatter diagram [2]. Among them, the global Moran I index generally reflects the distribution pattern of the entire study area; the global Geary C coefficient has a negative correlation with the global Moran I index and is also an index for analyzing global spatial autocorrelation; the global Getis index examines the clustering type of spatial distribution from a global perspective, that is, high-value agglomeration or low-value agglomeration; the local Moran I index reflects the distribution pattern state between each research area and surrounding areas, that is, whether there is an aggregation distribution; the Moran scatter plot reflects the relationship between a specific regional unit and its adjacent units Local spatial correlation mode [4]. These analysis methods involve the construction of a spatial weight matrix, the measurement and inspection of spatial autocorrelation, and the identification of spatial correlation. The following introduces the global Moran I test for testing whether each regional unit has global spatial autocorrelation and the test for the Moran scatterplot method of the spatial correlation pattern that the regional unit belongs to [5]. Suppose the observed value of the variable on each observed regional unit is known. In that case, the global Moran I index can be used to reflect the distribution of the index in the whole research area. The global Moran I test consists of two parts: constructing the spatial weight matrix and measuring spatial autocorrelation [8].

A binary symmetric spatial weight matrix $W_{nxn}$ is usually defined to express the spatial proximity relationship of a regional unit. The form is as follows:

$$W = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{bmatrix}$$
Among them, $w_{ij}$ represents the proximity relationship between regions $i$ and $j$, which can be measured according to distance or adjacency standards. The definition of the spatial weight matrix based on the distance criterion is as follows:

$$W_{ij} \begin{cases} 1, & \text{when the distance between } i \text{ and } j \text{ is less than } d \\ 0, & \text{other} \end{cases}$$

The definition of the spatial weight matrix based on the adjacency criterion is as follows:

$$W_{ij} \begin{cases} 1, & \text{when } i \text{ and } j \text{ are adjacent} \\ 0, & \text{other} \end{cases}$$

The Moran I index is defined as follows:

$$I = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{S^2 \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}}$$

Under the assumption of a normal distribution, the formulas for the expected value and variance of Moran I are as follows:

$$E(I) = \frac{1}{n-1}$$

$$VAR(I) = \frac{n^2 w_1 + nw_2 + 3w_0^2}{w_0^2(n^2 - 1)} - E^2(I)$$

Standardized statistics can be used to test whether there is a spatial autocorrelation relationship in an observation of a regional unit:

$$Z = \frac{I - E(I)}{\sqrt{VAR(I)}}$$

The value range of Moran I is between 0 and 1. When $|I| > 1.96$, the value of Moran I is positive, indicating that the data are positively correlated and the observed variables have a statistically significant space Cluster distribution. The closer the value of Moran I to 1, the stronger the positive spatial autocorrelation of the observed variable; when $|I| > 1.96$, the value of Moran I is negative, indicating that the data is negatively correlated, and the observed variable is meaningful. In addition, the closer the value of Moran I is to 1, the stronger the negative spatial autocorrelation of the observed variable is. On the contrary, when the value of Moran I is closer to 0, the data is randomly distributed and does not have spatial autocorrelation.
Table 1 Global Moran's I Test of GRP per capita

<table>
<thead>
<tr>
<th>Year</th>
<th>Moran Index(I)</th>
<th>E(I)</th>
<th>Var(I)</th>
<th>z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td>0.0362</td>
<td>-0.0333</td>
<td>0.0211</td>
<td>0.4783</td>
<td>0.3156</td>
</tr>
<tr>
<td>2003</td>
<td>0.0807</td>
<td>-0.0333</td>
<td>0.0211</td>
<td>0.7851</td>
<td>0.2148</td>
</tr>
<tr>
<td>2004</td>
<td>0.0990</td>
<td>-0.0333</td>
<td>0.0211</td>
<td>0.9105</td>
<td>0.1814</td>
</tr>
<tr>
<td>2005</td>
<td>0.0977</td>
<td>-0.0333</td>
<td>0.0211</td>
<td>0.9016</td>
<td>0.1841</td>
</tr>
<tr>
<td>2006</td>
<td>0.0881</td>
<td>-0.0333</td>
<td>0.0211</td>
<td>0.8359</td>
<td>0.2005</td>
</tr>
<tr>
<td>2007</td>
<td>0.1024</td>
<td>-0.0333</td>
<td>0.0211</td>
<td>0.9343</td>
<td>0.1762</td>
</tr>
</tbody>
</table>

The paper conducts a spatial analysis of the two variables, the financial-related ratio and the proportion of the added value of the financial industry in the regional GDP, respectively, according to the method introduced above. The adjacency standard determines the spatial weight matrix, as the elements of the spatial weight matrix corresponding to two adjacent regions are set to 1. Otherwise, they are 0. The global Moran I test is carried out on the two variables of the related financial ratio and the proportion of the added value of the financial industry in the regional GDP. For comparison, this paper also conducts a spatial analysis of the per capita GDP of 31 provinces and cities from 1997 to 2007. The empirical results are shown in Table 1.

Table 2 GLOBAL MORAN'S I TEST OF FINANCIAL INTERRELATION RATIO (FIR)

<table>
<thead>
<tr>
<th>Year</th>
<th>Moran Index(I)</th>
<th>E(I)</th>
<th>Var(I)</th>
<th>z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>1997</td>
<td>0.2969</td>
<td>-0.0333</td>
<td>0.0211</td>
<td>2.2726</td>
<td>0.0116</td>
</tr>
<tr>
<td>1998</td>
<td>0.2920</td>
<td>-0.0333</td>
<td>0.0211</td>
<td>2.2391</td>
<td>0.0125</td>
</tr>
<tr>
<td>1999</td>
<td>0.2892</td>
<td>-0.0333</td>
<td>0.0211</td>
<td>2.2198</td>
<td>0.0129</td>
</tr>
<tr>
<td>2000</td>
<td>0.2840</td>
<td>-0.0333</td>
<td>0.0211</td>
<td>2.1837</td>
<td>0.0146</td>
</tr>
<tr>
<td>2001</td>
<td>0.2814</td>
<td>-0.0333</td>
<td>0.0211</td>
<td>2.1658</td>
<td>0.0150</td>
</tr>
<tr>
<td>2002</td>
<td>0.2840</td>
<td>-0.0333</td>
<td>0.0211</td>
<td>2.1842</td>
<td>0.0145</td>
</tr>
<tr>
<td>2003</td>
<td>0.2957</td>
<td>-0.0333</td>
<td>0.0211</td>
<td>2.2644</td>
<td>0.0119</td>
</tr>
<tr>
<td>2004</td>
<td>0.2975</td>
<td>-0.0333</td>
<td>0.0211</td>
<td>2.2770</td>
<td>0.0113</td>
</tr>
<tr>
<td>2005</td>
<td>0.3074</td>
<td>-0.0333</td>
<td>0.0211</td>
<td>2.3450</td>
<td>0.0094</td>
</tr>
<tr>
<td>2006</td>
<td>0.3107</td>
<td>-0.0333</td>
<td>0.0211</td>
<td>2.3675</td>
<td>0.0089</td>
</tr>
<tr>
<td>2007</td>
<td>0.3070</td>
<td>-0.0333</td>
<td>0.0211</td>
<td>2.3423</td>
<td>0.0097</td>
</tr>
</tbody>
</table>

Table 3 GLOBAL MORAN'S I TEST OF THE SHARE OF VALUE-ADDED OF FINANCIAL INTERMEDIATION TO GRP

<table>
<thead>
<tr>
<th>Year</th>
<th>Moran Index(I)</th>
<th>E(I)</th>
<th>Var(I)</th>
<th>z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td>0.0170</td>
<td>-0.0333</td>
<td>0.0211</td>
<td>0.3462</td>
<td>0.3632</td>
</tr>
<tr>
<td>2003</td>
<td>0.0384</td>
<td>-0.0333</td>
<td>0.0211</td>
<td>0.4937</td>
<td>0.3121</td>
</tr>
<tr>
<td>2004</td>
<td>0.0337</td>
<td>-0.0333</td>
<td>0.0211</td>
<td>0.4612</td>
<td>0.3228</td>
</tr>
<tr>
<td>2005</td>
<td>0.0401</td>
<td>-0.0333</td>
<td>0.0211</td>
<td>0.5056</td>
<td>0.3050</td>
</tr>
<tr>
<td>2006</td>
<td>0.0367</td>
<td>-0.0333</td>
<td>0.0211</td>
<td>0.4821</td>
<td>0.3156</td>
</tr>
<tr>
<td>2007</td>
<td>0.0653</td>
<td>-0.0333</td>
<td>0.0211</td>
<td>0.6791</td>
<td>0.2483</td>
</tr>
</tbody>
</table>
From Table I that the Moran index I during the sample period is all positive. The Z values of the test are all greater than 1.96, and probability p values are all less than 0.05, indicating that at a confidence level of 0.05, The observed values of the per capita GDP of each province and city show a positive spatial correlation. The economic growth of each province and city shows a clustered distribution with statistical significance. Correspondingly, it can be seen from Table2 and Table 3 that even if the sample period starts from 2002, the spatial autocorrelation test results of the financial correlation ratio and the proportion of the added value of the financial industry in the regional GDP are still not significant, indicating that China The phenomenon of spatial aggregation distribution of financial development in various provinces and cities is not evident yet. However, it can be seen from Table 1 that the index and test values show an upward trend, and their significance level gradually increases, indicating that the spatial aggregation of financial development in various provinces and cities in China is getting higher and higher. From the spatial agglomeration of economic growth and the spatial agglomeration of financial development, economic agglomeration is far ahead of financial agglomeration.

4 CONCLUSION

This paper examines the phenomenon of regional financial agglomeration in China by using the global Moran I test for spatial statistical analysis, with reference to regional economic growth. It could be concluded that the economic growth of China's provinces and cities shows a statistically significant aggregation distribution, while the financial development is not statistically significant. However, the Moran index I and the test value Z basically show an increasing trend, and their significance levels gradually increase, indicating that the spatial agglomeration of financial development in China is getting higher and higher. This work indicates that the economic agglomeration is far ahead of financial agglomeration. (Sarkis, 2000)

REFERENCES

Business Strategy and Investment Efficiency

Xuetong Ding
Dxt0112@163.com
School of Economics and Management Beijing Jiaotong University Beijing, China

Abstract: The 2021 Central Economic Work Conference clearly stated that infrastructure investment should be moderately advanced. Efficient investment decisions can not only help enterprises grow, but also play an important role in promoting industrial technology innovation and creation. This paper obtains relevant data of A-share listed companies from 2010 to 2020 from CSMAR database, and uses empirical research methods to test the relationship between business strategy and investment efficiency. The conclusion is that business strategy has a significant impact on corporate investment efficiency. Compared with defenders, the impact on prospectors is greater. Such kind of companies’ investment efficiency is lower than that of defenders. The conclusions of this paper are still significant after the robustness test. Further analysis finds that the inefficiency of corporate investment caused by offensive strategies was mainly manifested as exacerbating overinvestment.

Keywords: Business Strategy, Investment Efficiency.

1 INTRODUCTION

The 2021 Central Economic Work Conference clearly stated that investment in infrastructure should be carried out moderately ahead of schedule, and the executive meeting of the State Council immediately made arrangements for expanding effective investment. For the country, expanding effective investment is an important starting point for stabilizing growth and promoting high-quality economic development. For enterprises, expanding effective investment actively will optimize the allocation of enterprise resources, which determines the development potential and competitiveness of enterprises.

Under the assumption of perfect competitive market, the marginal output of each economic project is equal. So the investment efficiency of enterprises only depends on the investment opportunities. Rational managers will invest limited funds in projects with positive NPV. However, in practice, this kind of market doesn’t exist. So, the investment decisions made by managers often deviate from the optimal investment level, either over-investment or under-investment, resulting in inefficient investment of enterprises. Investment behavior plays an important role in the daily operation and management of enterprises. Scholars’ exploration of the influencing factors of enterprise investment efficiency provides theoretical guidance for enterprises to improve investment efficiency. Taking together, the higher the investment efficiency of enterprises, the stronger the ability to activate corporate funds, the higher the profit margin, thereby enhancing the value of enterprises. However, in the actual business management activities, due to various factors, the phenomenon of low investment efficiency of
listed companies emerges in endlessly. Low investment efficiency not only will affect the value of the enterprise itself, but also may increase the risk of stock price crash, endangering the survival and development of enterprises. Therefore, how to improve the investment efficiency of enterprises has always been the focus of enterprises.

The business strategy determines a company’s future goals and long-term action plans for this purpose. Although the foothold of the business strategy is the future, it depends on the company’s current behavior and activities. The company’s strategy interacts with daily management behavior. The company relies on strategy to guide daily management behavior, and the realization of the company’s strategic goals is rooted in the continuous optimization of daily management activities. There are many ways to divide business strategy. The division method of Miles and Snow has been recognized by many scholars and has become the mainstream division method. The main reason is that this division method makes business strategy measurable, which is convenient for scholars to test its economic consequences through empirical research methods. Miles and Snow divide the company’s strategy into prospectors, analyzers and defenders. Prior papers show that different types of business strategies have different business characteristics and organizational structure. To maintain their leading role and leading position in the industry, prospectors tend to continue to develop new products and new markets. From the perspective of maintaining market share and ensuring competitiveness, defenders often focus on price, quality or service to develop competitive strategies. Based on this, we expect that prospectors may have more aggressive investment behavior, leading to overinvestment or under-investment, while defenders are more cautious about investment, inefficient investment behavior may be less. We expound the relationship between business strategy and corporate investment efficiency from two perspectives: information asymmetry and uncertainty face by companies. Due to the higher degree of information asymmetry and greater uncertainty, we expect that the non-investment efficiency of prospectors is aggravated.

Based on the above analysis, this paper obtains the relevant data of A-share listed companies from 2010 to 2020 from the CSMAR database. Then we use empirical research methods to test the relationship between business strategy and investment efficiency. The conclusion is consistent with our expectation that business strategy can influence investment efficiency and the impact of prospectors is more significant, and their investment efficiency is lower than that of defenders. Further analysis finds that prospectors have a lower investment efficiency mainly because the offensive strategy increases the over-investment of enterprises.

Compared with the existing papers, the main contributions of this paper are as follows: Firstly, the existing literature mainly studies the economic consequences of business strategy from the aspects of financial accounting behavior and consequences. This paper enriches the economic consequences of business strategy from the aspects of corporate investment efficiency. Secondly, the existing literature mainly explores the factors that affect the investment efficiency from three perspectives: external macro environment, internal features of the company and governance characteristics. The factors affecting the investment efficiency of enterprises are enriched from the perspective of internal features of the company. Thirdly, the practical significance of this paper is that since both investment behavior and business strategy are important components of internal financial management of enterprises, exploring the relationship between them plays an important role in optimizing internal financial management
of enterprises, which is conducive to guiding enterprises to choose a more suitable business strategy for themselves, and provides a way for enterprises to improve investment efficiency.

2 LITERATURE REVIEW

2.1 Factors influencing corporate investment efficiency

As an important part of enterprise management activities, investment behavior will not only affect the value of enterprises, but also affect the level of economic development (Xiong and Gu, 2022). Efficient investment decisions can not only help enterprises grow, but also play an important role in promoting industrial technology innovation. In addition, corporate investment activities are considered to be one of the three major activities to promote GDP growth, which plays a strong role in stimulating the rapid growth of China’s economy (Yan et al., 2021). Consequently, investment is vital for companies, industries and countries. However, due to the complexity and variability of the market environment, companies are not always able to make reasonable investment decisions, and inefficient investment behaviors are widespread (Shleifer and Vishny, 1988; Guariglia and Yang, 2016), leading to overinvestment or underinvestment or both. The research on the factors that influence investment efficiency of companies has never stopped. Current literature is mainly based on the these three aspects: First, the internal features of the company, including: accounting information comparability (Yuan and Rao, 2018), accounting information robustness (Zhang et al., 2017), property rights (Ma and Lin, 2021), etc.; Second, the governance layer characteristics, including: management ability (Yao et al., 2020), executive vertical concurrent (Yan et al., 2021), CEO discretion (Chen et al., 2020), etc.; Third, the external macro environment, including: government intervention (Wang et al., 2017), industrial policy (Lin and Zhang, 2022), product market competition (Wang et al., 2019), etc. These studies provide a theoretical basis for enterprises to seek more efficient investment. However, most of these studies ignore the economic consequences of the investment efficiency of enterprises.

2.2 Economic Consequences of Business Strategy

Although a company's strategic choice cannot be simply defined as a certain type of strategy, it’s conducive to studying the economic consequences of business strategy by dividing different strategic types and classifying each company into the most suitable strategic type. The mainstream types of business strategy in management are: overall cost leadership and differentiation (Porter, 1980); exploratory and exploitative (March, 1991); intimate customer, efficient operation and product leading (Treacy et al., 1995); Miles and Snow (1978, 2003) divided the strategy into prospectors, analyzers and defenders, according to the radical degree of the strategy adopted by the company. While other classification methods’ data can only be obtained through interviews or surveys, making it difficult for data acquisition, Miles and Snow’s (1978, 2003) classification method can cover the mainstream classification and can be measured by company's financial data, which is convenient for scholars to conduct empirical analysis. Therefore, Miles and Snow’s (1978, 2003) classification method is widely used.

Companies that have different business strategies distinct from each other in operating and organizational structure. Generally, prospectors have more operating projects, wider business
scope, and more complex internal structure, which provides chances and space for management manipulation like earnings management (Sun et al., 2016), excessive investment (Wang et al., 2016), company violations (Meng et al., 2018), etc. Defenders often have a narrower business scope and a more centralized organizational structure, which is more conducive to the implementation of regulations. It can effectively regulate the behavior of management and reduce the space for management to seek personal interests. The above differences between two types of strategies will finally reflect in and affect the enterprise value. Prospectors tend to invest a lot of money in technology research and development, product creation and market development. While such activities can enhance the growth of companies, they can also increase risks faced by companies now or soon: if successful, the company's performance will increase significantly and achieve value growth; if unsuccessful, the company will suffer huge losses, exacerbating the risk of corporate stock crashes and negatively affecting corporate value. On the contrary, defenders mainly focus on improving the quality of developed products and services and retaining their market share. Therefore, compared with prospectors, defenders have better performance stability, at the same time, they may also lose future growth and may have a negative impact on the future value of the companies.

3 HYPOTHESIS DEVELOPMENT

3.1 Information Asymmetry

From the perspective of information asymmetry, the degree of information asymmetry of prospectors is higher than that of defenders. On the one hand, Liu et al. (2015) believe that when information asymmetry's level is high, the effectiveness of supervision by potential external investors on management will be greatly reduced due to their inability to obtain sufficient information, which provides an opportunity for management to implement opportunistic investment behavior. Overinvestment behavior may occur. On the other hand, due to the existence of information gap, the party who grasps more real and effective information may benefit from its own information advantages by obtaining additional benefits. When making investment decisions, enterprises are in the inferior side of the information. It means they cannot fully obtain the information of the investor or the invested unit. Meanwhile, the investor or the invested unit are in the information advantage side. They tend to hide the information that is not conducive to the investment project, which will affect the analysis and prediction of the investment project by the enterprise managers, resulting in the wrong estimation of the investment return of the project by the enterprise. Finally, the enterprise cannot identify the project with the highest investment return due to the wrong estimation. Missing the optimal investment project reduces the investment efficiency of the enterprise.

3.2 Uncertainty faced by Enterprises

From the perspective of the uncertainty faced by enterprises, prospectors face greater uncertainty. Compared with defenders, prospectors will invest a lot of funds for continuous market development, research and development of new products, etc. These behaviors are large in number and amount and frequent in changes, making the company's operating environment more unstable. The performance risks faced by companies are so high that they encounter a higher degree of uncertainty. If the investment is successful, it can quickly improve
the enterprise value and competitiveness. If the investment fails, their huge funds may not be recovered or liquidation, exacerbating the risk of corporate share price crash. This uncertainty increases the difficulty of predicting the best level of investment by enterprises, which makes the prediction results deviate greatly from the actual situation. So, it’s likely to mislead decision makers to make investment decisions that are not conducive to enterprises and reduce the investment efficiency of enterprises.

Taking together, we posit the following hypothesis:

H1: The investment efficiency of prospectors is lower than that of defenders.

4 SAMPLE, VARIABLES AND RESEARCH DESIGN

4.1 Sample Selection

The data used in this study are collected from the China Stock Market and Accounting Research (CSMAR) database. The sample selection begins with all A-share listed companies from 2010 to 2020, then, following prior paper, eliminates financial firms and other firms that have missing data or negative equity. Finally, the sample contains 18124 firm-year observations. To minimize the effect of outliers, all continuous variables are winsorized at 1st and 99th percentiles.

4.2 Variables

4.2.1 Business Strategy

Based on Bentley et al. (2013), the strategy is composed of six measures, each of which is measured by the rolling prior five-year average. Then sort all variables by industry and year. For the first five variables, the highest quintiles are given a score of 5, the second highest quintiles are given a score of 4, and so on, while the lowest quintiles are given a score of 1. For the last variable, which is reversed-scored, the lowest (highest) quintile are given a score of 5 (1). Finally, we get each company’s scores (Strategy) ranging from 6 to 30 by adding the six measures per company-year. In this paper, we consider the following strict definitions of strategy-types: defender-type (6–18); prospector-type (18–30). The standard of classification (the score of 18) is the annual median of all company’s scores. Refer to Table 1 for details related to the measurement of business strategy.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Company’s propensity to search for new products (STR1)</td>
<td>Ratio of R&amp;D expenditure to sales.</td>
</tr>
<tr>
<td>(2) Company’s ability to produce and distribute products and services efficiently (STR2)</td>
<td>Ratio of the number of employees to sales.</td>
</tr>
<tr>
<td>(3) Company’s historical growth or investment opportunities (STR3)</td>
<td>One-year percentage change in total sales.</td>
</tr>
</tbody>
</table>
### 4.2.2 Prospectors

An indicator variable equals to 1 if a company belongs to prospector-type (the calculated score of the company is greater than the annual median (18)), and otherwise 0.

### 4.2.3 Investment Efficiency

This paper draws on Richardson (2006), Wang (2011), Zhang (2014) and other papers’ estimation methods of the investment efficiency, and establishes a specific model (1) as follows:

\[
\text{Invest}_{t} = \alpha_{0} + \alpha_{1}\text{Growth}_{t-1} + \alpha_{2}\text{Cash}_{t-1} + \alpha_{3}\text{Size}_{t-1} \\
+ \alpha_{4}\text{Age}_{t-1} + \alpha_{5}\text{Lev}_{t-1} + \alpha_{6}\text{R}_t + \alpha_{7}\text{Inv}_{t-1} \\
+ \sum \text{Industry} + \sum \text{Year} + \epsilon_{t}
\]

(1)

Often, we use the absolute value of the residual estimated by this model (\(INV\)) to measure the investment efficiency of a company. Refer to Table 2 for details related to the specific definition and measurement of each variable in this model.

Control variables: The control variables are chosen based on the literature (e.g., Lv and Zhang 2011; Cheng et al. 2012), including Growth, Cash, Size, Age, Le, R, Top1, Loss, and Roa. The definitions of these variables are shown in Table 2.

### 4.3 Research Design

To examine the influence of business strategy on investment efficiency, this paper establishes the regression model (2) as follows:

\[
\text{INV}_{it} = \beta_{0} + \beta_{1}\text{Strategy}_{it} + \beta_{2}\sum \text{Controls}_{it} \\
+ \sum \text{Industry} + \sum \text{Year} + \epsilon_{it}
\]

(2)

According to H1, we expect \(\beta_{1}\) to be positive and statistically significant. It means that the more aggressive the business strategy, the greater the value of \(\text{INV}\), the higher the degree of inefficient investment, the lower the efficiency of corporate investment.

In order to show the impact of different business strategy types on corporate investment efficiency more clearly, this paper replace Strategy with Prospectors in model (2). The regression model (3) is as follows:

---

**Variables** | **Measurement**
--- | ---
(4) Company’s focus on exploiting new products and services. \((\text{STRA}_{1})\) | Ratio of selling and administrative expenses to sales.
(5) Company’s organizational stability \((\text{STRA}_{3})\) | Standard deviation of the total number of employees.
(6) Company’s commitment to technological efficiency \((\text{STRA}_{6})\) | Capital intensity which is measured as net PPE scaled by total assets.

*All variables are computed over a rolling prior five-year average.*
\[ \text{INV}_u = \beta_0 + \beta_1 \text{Prospectors}_u + \beta_2 \sum \text{Controls}_u \]
\[ + \sum \text{Industry} + \sum \text{Year} + \epsilon_u \quad (3) \]

\( \theta_i \) in model (3) quantifies the difference between the impact of the two strategy types on company’s investment efficiency. According to H1, we expect \( \theta_i \) to be positive and statistically significant, indicating that after controlling other variables, the impact of prospectors on corporate investment efficiency is greater than defenders. The main variables and descriptions in this model are shown in Table 2.

**TABLE II. VARIABLE DEFINITIONS**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Independent variable</strong></td>
<td></td>
</tr>
<tr>
<td><strong>INV</strong></td>
<td>The absolute value of the residual estimated by model (1). The greater the value, the higher the degree of inefficient investment, that is, a company with a greater INV has a lower investment efficiency.</td>
</tr>
<tr>
<td><strong>Dependent Variables</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Strategy</strong></td>
<td>A measure of business strategy, ranging from 6-30 scores. The greater the value, the more aggressive the company’s strategy.</td>
</tr>
<tr>
<td><strong>Prospectors</strong></td>
<td>An indicator variable equals to 1 if a company belongs to prospector-type (the calculated score of the company is greater than the annual median (18)), and otherwise 0.</td>
</tr>
<tr>
<td><strong>Control variables</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Growth</strong></td>
<td>Using tobin Q as an proxy.</td>
</tr>
<tr>
<td><strong>Cash</strong></td>
<td>Cash and net short-term investments divided by total assets.</td>
</tr>
<tr>
<td><strong>Size</strong></td>
<td>The natural logarithm of total assets.</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td>The natural logarithm of current year minus listed year plus 1.</td>
</tr>
<tr>
<td><strong>Lev</strong></td>
<td>Total liabilities divided by total assets.</td>
</tr>
<tr>
<td><strong>R</strong></td>
<td>Annual stock return minus comprehensive annual market return.</td>
</tr>
<tr>
<td><strong>Top1</strong></td>
<td>The proportion of the largest shareholder in the total number of shares.</td>
</tr>
<tr>
<td><strong>Loss</strong></td>
<td>An indicator variable equals to 1 if a company has a negative net profit, and otherwise 0.</td>
</tr>
<tr>
<td><strong>Roa</strong></td>
<td>The net income before extraordinary items scaled by total assets.</td>
</tr>
</tbody>
</table>

**5 EMPIRICAL RESULTS**

**5.1 Descriptive Statistics**

Table 3 reports descriptive statistics for our main variables. The gap between the maximum and minimum of business strategy (Strategy) is large, indicating that there is a big difference in the strategic choices between listed companies. The mean and standard deviation of Strategy is 17.807, and 4.175, indicating that the sample distribution is basically reasonable—defender-type companies and prospector-type companies coexist. The mean and standard
deviation of investment efficiency ($INV$) is 0.040 and 0.058, which is not much different from the statistical value of existing research. Other variables are all within a reasonable range.

**TABLE III. DESCRIPTIVE STATISTICS**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Obs</th>
<th>Mean</th>
<th>Std</th>
<th>P50</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>INV</td>
<td>18124</td>
<td>0.040</td>
<td>0.058</td>
<td>0.024</td>
<td>0</td>
<td>0.960</td>
</tr>
<tr>
<td>Strategy</td>
<td>18124</td>
<td>17.807</td>
<td>4.175</td>
<td>18</td>
<td>6</td>
<td>30</td>
</tr>
<tr>
<td>Prospector</td>
<td>18124</td>
<td>0.436</td>
<td>0.496</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Growth</td>
<td>18124</td>
<td>2.143</td>
<td>2.597</td>
<td>1.596</td>
<td>0.153</td>
<td>122.189</td>
</tr>
<tr>
<td>Cash</td>
<td>18124</td>
<td>0.040</td>
<td>0.058</td>
<td>0.024</td>
<td>0</td>
<td>0.960</td>
</tr>
<tr>
<td>Size</td>
<td>18124</td>
<td>22.440</td>
<td>1.290</td>
<td>22.294</td>
<td>17.641</td>
<td>28.257</td>
</tr>
<tr>
<td>Age</td>
<td>18124</td>
<td>2.588</td>
<td>0.445</td>
<td>2.639</td>
<td>1.609</td>
<td>3.434</td>
</tr>
<tr>
<td>Lev</td>
<td>18124</td>
<td>0.472</td>
<td>0.210</td>
<td>0.473</td>
<td>0.007</td>
<td>3.919</td>
</tr>
<tr>
<td>R</td>
<td>18124</td>
<td>0.017</td>
<td>0.449</td>
<td>-0.078</td>
<td>-1.057</td>
<td>7.120</td>
</tr>
<tr>
<td>Top1</td>
<td>18124</td>
<td>33.107</td>
<td>14.797</td>
<td>30.600</td>
<td>0.290</td>
<td>89.990</td>
</tr>
<tr>
<td>Loss</td>
<td>18124</td>
<td>0.127</td>
<td>0.333</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Roa</td>
<td>18124</td>
<td>0.024</td>
<td>0.110</td>
<td>0.029</td>
<td>-4.946</td>
<td>0.786</td>
</tr>
</tbody>
</table>

**5.2 Empirical Results**

This paper uses model (2) to test the relationship between business strategy and investment efficiency. Table 4 presents the main results. In column (1), the coefficient on $Strategy$ is significantly positive (t-stat = 4.99), indicating that there is a positive association between business strategy and $INV$. It means business strategy has a significant impact on corporate investment efficiency.

To identify the impact comes from which type of strategy, we use model (3) to test which type of business strategy will have a more significant effect on investment efficiency. The results are shown in Column (2) of Table 4. It can be seen from the table that the coefficient on $Prospectors$ is 0.00441 and is significantly positive at 1 percent level, indicating that under the premise that other variables remain unchanged, prospectors’ impact on $INV$ is greater than defenders. It means other variables remain unchanged, the value of companies adopting prospector-type strategies is greater than that of companies adopting defender-type strategies, so the investment efficiency of companies adopting prospector-type strategies is lower. The result is consistent with H1.

**TABLE IV. BUSINESS STRATEGY AND INVESTMENT EFFICIENCY**

<table>
<thead>
<tr>
<th>Variables</th>
<th>INV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>Strategy</td>
<td>0.977***</td>
</tr>
<tr>
<td>Prospector</td>
<td></td>
</tr>
</tbody>
</table>
6  ROBUSTNESS TESTS

6.1 Endogeneity Problem

If the independent variable business strategy (Strategy) is related to the random error term (\( \varepsilon_a \)), the model has an endogenous problem. To alleviate this concern, we introduce STRA_mean (the 'industry-annual' strategic heterogeneity mean), which was first proposed by Wang et al. (2016). The regression results are shown in Table 5 column (1) and (2). The first stage regression results show that the instrumental variable (STRA_mean) is significantly positively correlated with the core independent variable (Strategy), indicating that the endogenous problem will not affect the above regression results, and the conclusion of this paper is robust.

6.2 Change Interested Variable’s Measurement

Referring to the practice of Ye et al. (2014), we replace the measure of STRA_t. In specific, we use 'the change value of intangible assets' as an alternative for 'R&D expenditure'. The
measure of $\text{STRA}_t$ becomes 'Ratio of the change value of intangible assets to sales.', and other variables' measurement remain unchanged. Then a new business strategy index is calculated and is used to re-examine the relationship between business strategy and investment efficiency. The regression results are shown in Table 5 column (3) and (4), which are consistent with our main conclusion.

### TABLE V. ROBUSTNESS TESTS

<table>
<thead>
<tr>
<th>Variables</th>
<th>Strategy</th>
<th>INV (1)</th>
<th>INV (2)</th>
<th>INV (3)</th>
<th>INV (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strategy</td>
<td>0.977***</td>
<td>0.00104***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(9.42)</td>
<td>(6.79)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prospects</td>
<td>0.00297*</td>
<td></td>
<td></td>
<td>0.00639***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.80)</td>
<td></td>
<td></td>
<td>(5.91)</td>
<td></td>
</tr>
<tr>
<td>Growth</td>
<td>0.0577***</td>
<td>0.000340</td>
<td>0.000463**</td>
<td>0.000504***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.43)</td>
<td>(1.23)</td>
<td>(2.48)</td>
<td>(2.61)</td>
<td></td>
</tr>
<tr>
<td>Cash</td>
<td>0.898**</td>
<td>-0.0112*</td>
<td>-0.000574</td>
<td>-0.000584</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.69)</td>
<td>(-1.89)</td>
<td>(-0.12)</td>
<td>(-0.13)</td>
<td></td>
</tr>
<tr>
<td>Size</td>
<td>1.015**</td>
<td>-0.00133</td>
<td>0.000749</td>
<td>0.000902</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(13.23)</td>
<td>(-1.50)</td>
<td>(1.21)</td>
<td>(1.44)</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-2.371***</td>
<td>-0.00339</td>
<td>-0.00897***</td>
<td>-0.00950***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-14.46)</td>
<td>(-0.94)</td>
<td>(-6.27)</td>
<td>(-6.68)</td>
<td></td>
</tr>
<tr>
<td>Lev</td>
<td>0.0523</td>
<td>0.00836*</td>
<td>0.00474</td>
<td>0.00453</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.18)</td>
<td>(1.87)</td>
<td>(1.27)</td>
<td>(1.21)</td>
<td></td>
</tr>
<tr>
<td>$B^*$</td>
<td>-0.0596</td>
<td>0.00735***</td>
<td>0.00706***</td>
<td>0.00713***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-4.11)</td>
<td>(4.65)</td>
<td>(4.51)</td>
<td>(4.54)</td>
<td></td>
</tr>
<tr>
<td>Top1</td>
<td>0.00249</td>
<td>0.00000288</td>
<td>-0.0000596</td>
<td>-0.0000704</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.48)</td>
<td>(0.04)</td>
<td>(-1.29)</td>
<td>(-1.53)</td>
<td></td>
</tr>
<tr>
<td>Loss</td>
<td>-0.221***</td>
<td>-0.00166</td>
<td>-0.00396***</td>
<td>-0.00399***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-3.28)</td>
<td>(-1.20)</td>
<td>(-2.78)</td>
<td>(-2.79)</td>
<td></td>
</tr>
<tr>
<td>Roa</td>
<td>-0.526*</td>
<td>0.00666</td>
<td>0.00969*</td>
<td>0.00972*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-1.77)</td>
<td>(1.43)</td>
<td>(1.79)</td>
<td>(1.78)</td>
<td></td>
</tr>
<tr>
<td>cons</td>
<td>-15.95***</td>
<td>0.0227</td>
<td>0.0290**</td>
<td>0.0432***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-6.58)</td>
<td>(0.81)</td>
<td>(2.12)</td>
<td>(3.29)</td>
<td></td>
</tr>
<tr>
<td>Year/Industry</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>18124</td>
<td>18124</td>
<td>18051</td>
<td>18051</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.1136</td>
<td>0.0314</td>
<td>0.0460</td>
<td>0.0441</td>
<td></td>
</tr>
</tbody>
</table>

a. Standard errors are clustered at the firm level.
b. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

### 7 ADDITIONAL ANALYSIS

As known, inefficient investment includes overinvestment and underinvestment. To explore the specific performance of inefficient investment of prospectors, we then conduct groping regression and the grouping criteria are the residual in model (1). Specifically, the sample with residuals greater than zero in model (1) is the overinvestment group and the samples with residuals less than zero in model (1) are the underinvestment group. The results are shown in
Table 6. In the overinvestment group, the coefficients are significant at the 1%, indicating that the inefficient investment of prospectors is mainly reflected in overinvestment.

<table>
<thead>
<tr>
<th>TABLE VI. GROUPING REGRESSION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variables</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Strategy</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Prospector</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Growth</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Cash</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Size</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Age</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Lev</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>R</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Top1</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Loss</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Roi</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>cons</td>
</tr>
<tr>
<td>Year/Industry</td>
</tr>
<tr>
<td>N</td>
</tr>
<tr>
<td>R²</td>
</tr>
</tbody>
</table>

a. Standard errors are clustered at the firm level.
b. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

8 CONCLUSION

This paper finds that prospectors usually aggravates the inefficient investment behavior which results in lower investment efficiency. Further analysis finds that the offensive strategy leads to lower investment efficiency, mainly due to the offensive strategy increases overinvestment.
The guiding significance of the research conclusion of this paper lies in: First, for managers, it is necessary to realize that the role of business strategy has two sides. On the one hand, prospector-type strategy can speed up the development of the company to a certain extent, and enhance the value and the development momentum of the enterprise. On the other hand, prospector-type strategy often brings problems like excessive financial leverage and too wide business scope, which may have a negative effect on the company's future performance and inevitably increase the company's risks. Second, for regulators, it is necessary to strengthen the supervision of the company and enhance the external governance environment. Restricting overly aggressive investment decisions of listed companies can reduce the probability of stock price crashes. Such kind of restrictions play an important role in preventing systemic financial risks and maintaining the order and stability of the capital market. Nowadays, Chinese enterprises are still facing the problem of low investment efficiency. With the advancement of supply-side structural reform, enterprises need to continuously improve the efficiency of capital use in order to adapt to the national policy trend, which can alleviate this problem to a certain extent and promote economic growth. Enterprises should not only always pay attention to the possible changes in corporate investment behavior, but also carefully make strategic change decisions—strategy should be in line with the actual situation of the enterprise.

REFERENCES

An Empirical Study of ESG Performance and Corporate Investment Efficiency
-- Moderating Effect of External Pressure

Chunxia Fu\textsuperscript{1a}, Zhiguang Guo\textsuperscript{1b}, and Luorao Yang\textsuperscript{2c}

\textsuperscript{a}431067793@qq.com, \textsuperscript{b}369346618@qq.com, \textsuperscript{c}1735474591@qq.com

\textsuperscript{1}College of Business and Tourism, Sichuan Agricultural University, Dujiangyan, Chengdu, China
\textsuperscript{2}Department of Business Management, Sichuan Agricultural University, Chengdu, China

Abstract: This paper develops a multiple linear regression model to empirically test the effect of corporate ESG performance on investment efficiency and using the nature of equity and media attention as the entry point to deeply explore the moderating effect of external pressure on ESG performance and investment efficiency, selecting Shanghai and Shenzhen A-share main board listed companies from 2012 to 2020 as the research object. The findings indicate that ESG performance significantly suppresses inefficient investment, i.e., better ESG performance is correlated with higher investment efficiency; however, external pressures on companies, such as the nature of equity and media attention, mitigate the effect of ESG performance on investment efficiency. The empirical results of this paper further highlight the existence of specificities in the Chinese capital market, while providing important theoretical insights for the further implementation of ESG practices in China.

Keywords: ESG performance; nature of equity; media attention; investment efficiency; multivariable linear regression model.

1 INTRODUCTION

The abbreviation ESG stands for environment, social, and governance, and it refers to an investment idea and comprehensive evaluation criteria that focuses on environmental pollution management, social duty fulfillment, and corporate governance performance. The 18th national Congress of CCPC established construction of ecological civilization as a strategic aim for the first time. As a result of this potential, the green economy and sustainability concept have grown into an unavoidable trend now and for a longer period, the ESG concept with sustainable development as its core has gradually been noticed and accepted.

In recent years, several theoretical and practical investigations on ESG have been conducted in academia. Better ESG performance, according to well-researched studies on ESG behavior at the business level, may considerably cut the cost of company financing (Qiu and Yin 2019)\textsuperscript{3}, raise the book value and market value of corporations (Wang and Yang 2022)\textsuperscript{8}, and promote corporate innovation (Zhang et al. 2020)\textsuperscript{13}. In addition, companies with superior ESG performance have greater excess returns (Zhou et al. 2020)\textsuperscript{14}, and tend to exhibit better performance levels (Yuan and Xiong 2021)\textsuperscript{12}. Furthermore, the impact of external pressure on corporate development should not be overlooked. The government can exercise control over...
companies through legal penalties and policy restrictions, the media may exert pressure on companies to commence ESG behavior through news coverage, and third-party stakeholders can influence corporate investment decisions (Song et al. 2019)[6].

In summary, domestic scholars have rarely focused on this area, and there is almost no discussion on the mechanism of external pressure in related literature, leaving greater space for discussion. Thus, this paper selects Shanghai and Shenzhen A-share main board listed companies as the research object, develops a multiple linear regression model, discusses the effect of corporate ESG performance on investment efficiency, investigates the difference between the two relationships as external pressure changes, and provides the theoretical foundation for the subsequent development of ESG practice in corporations.

2 RESEARCH HYPOTHESIS

2.1 ESG performance and Investment Efficiency

According to the perspective of stakeholder theory and resource dependence theory, good ESG performance can partially satisfy the demands of various stakeholders, increase their chances of gaining stakeholder support, which makes it more likely to obtain relevant external resources, and ultimately improve the effectiveness of their investment decisions (Anwar and Malik 2020)[1]. As a result, hypothesis 1 is offered.

H1: Corporate investment efficiency can be increased with strong ESG performance.

2.2 Moderating Effect of External Pressure

According to the viewpoint of organizational legitimacy, it is crucial for a company's survival and growth that its behavior conforms to the expectations of its stakeholders. Researcher have found that public opinion monitoring and government regulation can significantly improve firms' environmental information disclosure (Shen and Feng 2012)[5]. Therefore, this paper proposes hypotheses in turn.

Government departments and other regulators have passed a series of laws, regulations, and policy documents to encourage listed corporations to incorporate ESG practices into their development strategies, improve corporate ESG information disclosure, enhance their ESG performance, and guide investors to practice ESG investment concepts, thus resulting in external pressure that can influence corporations' investment decisions to some extent. Furthermore, because of the nature of the equity of the corporation, this form of external pressure will be far more binding on SOEs than non-SOEs. As a result, SOEs may emphasize investing in environmental and social responsibility, leading to investment inefficiencies. As a result, hypothesis 2 is offered.

H2: The relationship between ESG performance and investment efficiency is moderated by the nature of equity.

As information technology advances, the media may play an important external monitoring role in the stock market (Walter et al. 2008)[7]. Firstly, media coverage can improve stakeholders' comprehension of corporate ESG performance, minimize information asymmetry, lower investment risk, and increase its investment efficiency; then, for those companies with
insufficient ESG performance, the media may focus on reporting their negative information, and under such public pressure, companies may over-invest in environmental and social responsibility activities. Such passive ESG practices may consume their resources, and media attention may play a reverse moderating role. As a result, hypothesis 3 is offered.

H3: The relationship between ESG performance and investment efficiency is moderated by media attention.

3 RESEARCH DESIGN

3.1 Sample Selection and Data Sources

This paper selects listed businesses on the main board of Shanghai and Shenzhen A-shares from 2012 to 2020 as an initial sample and screen the initial sample using the criteria listed below: (1) exclude the sample of the financial industry; (2) exclude the sample of ST, *ST and PT; (3) exclude the sample of assets and liabilities ratio over 1; (4) exclude the sample listed after 2020; (5) exclude the variable samples with more missing values. Finally, we obtained 1076 valid samples to construct a balanced panel database with 9684 observations. The data are obtained from the Wind database (Wind), the China Listed Companies Financial News Database (CFND), and the CSMAR database (CSMAR). During the study, this paper's data processing and statistical analysis were carried out using Excel 2013 and Stata 17.

3.2 Variable Description

3.2.1 Explained Variable

Corporate investment efficiency (Inv). Corporate investment efficiency reflects whether firms can fully use investment opportunities for value creation. Referring to Richardson (2006) and Xu (2014), the absolute value of the residuals generated from OLS regressions of the model (1) by industry and year is used to measure corporate investment efficiency in this paper.

\[
Inv_{it} = a_0 + a_1 Growth_{it-1} + a_2 Leve_{it-1} + a_3 Cash_{it-1} + a_4 Age_{it-1} + a_5 Size_{it-1} + \sum Industry + \sum Year + \varepsilon
\]  

3.2.2 Explanatory Variable

ESG performance (ESG). ESG Ratings establish nine levels according to the ESG level of the corporation, and this paper uses the nine-point system to rate the ESG performance of the company: 9 points for AAA, 8 points for AA, 7 points for A, 6 points for BBB, 5 points for BB, 4 points for B, 3 points for CCC, 2 points for CC, and 1 point for C. The higher the score, the better the ESG performance.

3.2.3 Adjustment Variables

Nature of equity (Soe). In this study, we use dummy variables 0-1 to measure the nature of equity. If the corporation's true controller is the State-owned Assets Supervision, Administration Commission or government department at all levels, it has been determined to be a state-owned corporation and takes the value of 1. Other non-state-owned corporations take a value of 0.
Media attention (Media). This paper adopts newspaper financial news from CFND as the data source for media attention and utilizes the "1 + number of media coverage" natural logarithm to judge the degree of media attention of corporations.

3.2.4 Control Variables

In this study, which references earlier pertinent studies (Yao et al. 2020[11]; Wang et al. 2021[9]), the following control variables are selected: age of firm listing (Age), firm size (Size), firm growth (Growth), total net asset margin (Roa), gearing ratio (Lev), management remuneration (Sala), equity concentration (Hold), and the percentage of independent directors (Dboard). Meantime, year and industry dummy variables are introduced to adjust for time and industry impacts.

3.3 Multiple Linear Regression Models

In this study, we set models (2) to (4) for testing hypotheses H1 to H3, with the following model settings:

\[
\text{Inv}_{i,t} = \beta_0 + \beta_1 \text{ESG}_{i,t} + \sum \beta_j \text{C}_{i,t} + \text{Industry} + \text{Year} + \epsilon_{i,t} \\
\text{(2)}
\]

\[
\text{Inv}_{i,t} = \beta_0 + \beta_1 \text{ESG}_{i,t} + \beta_2 \text{State}_{i,t} + \beta_3 \text{ESG}_{i,t} \times \text{State}_{i,t} + \sum \beta_j \text{C}_{i,t} + \text{Industry} + \text{Year} + \epsilon_{i,t} \\
\text{(3)}
\]

\[
\text{Inv}_{i,t} = \beta_0 + \beta_1 \text{ESG}_{i,t} + \beta_2 \text{Media}_{i,t} + \beta_3 \text{ESG}_{i,t} \times \text{Media}_{i,t} + \sum \beta_j \text{C}_{i,t} + \text{Industry} + \text{Year} + \epsilon_{i,t} \\
\text{(4)}
\]

Model (2) investigates the direct impact of ESG and Inv. Model (3) tests the role of equity nature in modulating ESG performance and investment efficiency. Model (4) tests the role of media attention in modulating ESG performance and investment efficiency.

4 EMPIRICAL RESULTS AND ANALYSIS

4.1 Descriptive Statistics and Correlation Analysis

In Table 1, the sample corporations generally have inefficiency in investment with significant differences. The mean value of ESG is 4.250, showing that the sample corporate average ESG grades vary from CCC to B, indicating that ESG performance still needs to be improved.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inv</td>
<td>0.038</td>
<td>0.068</td>
<td>0</td>
<td>2.885</td>
</tr>
<tr>
<td>ESG</td>
<td>4.250</td>
<td>1.096</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>Age</td>
<td>2.624</td>
<td>0.495</td>
<td>1.099</td>
<td>3.332</td>
</tr>
<tr>
<td>Size</td>
<td>22.75</td>
<td>1.325</td>
<td>20.21</td>
<td>26.65</td>
</tr>
<tr>
<td>Growth</td>
<td>0.092</td>
<td>0.252</td>
<td>-0.522</td>
<td>1.170</td>
</tr>
<tr>
<td>Roa</td>
<td>0.037</td>
<td>0.045</td>
<td>-0.112</td>
<td>0.187</td>
</tr>
</tbody>
</table>
Table 2 shows the results of a bivariate inter-variate Pearson analysis performed on the variables listed above. The correlation coefficients show that ESG and Inv are highly negatively associated, with a correlation coefficient of -0.078, implying that ESG performance has a favorable influence on corporate investment efficiency. Furthermore, the mean value of VIF across variables is 1.40 and its maximum value is 2.25, which is substantially lower than the reference standard value of 5.

<table>
<thead>
<tr>
<th></th>
<th>Inv</th>
<th>ESG</th>
<th>Age</th>
<th>Size</th>
<th>Growth</th>
<th>Roa</th>
<th>Lev</th>
<th>Sala</th>
<th>Hold</th>
<th>Dboard</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inv</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ESG</td>
<td>-0.078 ***</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-0.076 ***</td>
<td>0.018 #</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size</td>
<td>-0.089 ***</td>
<td>0.286 ***</td>
<td>0.219 ***</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Growth</td>
<td>0.092 ***</td>
<td>0.033 ***</td>
<td>0.098 ***</td>
<td>-0.053 ***</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roa</td>
<td>0.044 ***</td>
<td>0.147 ***</td>
<td>0.092 ***</td>
<td>-0.024 ***</td>
<td>0.213 ***</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lev</td>
<td>-0.063 ***</td>
<td>0.027 ***</td>
<td>0.204 ***</td>
<td>0.526 ***</td>
<td>0.043 ***</td>
<td>-0.386 ***</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sala</td>
<td>-0.028 ***</td>
<td>0.248 ***</td>
<td>0.131 ***</td>
<td>0.532 ***</td>
<td>0.093 ***</td>
<td>0.192 ***</td>
<td>0.179 ***</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hold</td>
<td>0.026 **</td>
<td>0.074 ***</td>
<td>0.038 ***</td>
<td>0.243 ***</td>
<td>-0.040 ***</td>
<td>0.090 ***</td>
<td>0.106 ***</td>
<td>0.052 ***</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Dboard</td>
<td>-0.006 ***</td>
<td>0.126 ***</td>
<td>-0.005 ***</td>
<td>0.113 ***</td>
<td>0.005 ***</td>
<td>-0.026 ***</td>
<td>0.035 ***</td>
<td>0.043 ***</td>
<td>0.041 ***</td>
<td>1</td>
</tr>
</tbody>
</table>

### 4.2 Baseline Regression Analysis

In order to better address the issue of omitted variables, improve the estimation accuracy of the model, and represent more information on the form of individual dynamics, this paper uses balanced panel data for a total of nine years from 2012 to 2020. The Hausman test was first performed before the regression analysis, and the result showed a p-value of 0. Therefore, it was decided that the panel regression of the data in this study would adopt a fixed effects model. In
addition, the regressions were clustered at the firm level for standard errors to prevent the impact of aggregation effects on standard errors at the firm level.

Table 3 displays the results of the empirical model's multiple linear regressions. After controlling for industry and year fixed effects, column (1) can be seen that ESG is negatively correlated with Inv, implying that there is a boosting effect of ESG on Inv. Further, column (2) presents the results after adjusting for all control variables, and the ESG regression coefficient is -0.003, implying that corporate ESG performance successfully suppresses its inefficient investment and positively enhances its investment efficiency. As a result, the empirical data support hypothesis H1 of this research.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ESG</td>
<td>-0.004***</td>
<td>-0.003***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.043***</td>
<td>0.080***</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.029)</td>
</tr>
<tr>
<td>Controls</td>
<td>NO</td>
<td>Yes</td>
</tr>
<tr>
<td>Industry</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>N</td>
<td>9684</td>
<td>9684</td>
</tr>
<tr>
<td>R²</td>
<td>0.054</td>
<td>0.067</td>
</tr>
<tr>
<td>Adj.R²</td>
<td>0.047</td>
<td>0.059</td>
</tr>
</tbody>
</table>

The figures in parenthesis represent standard deviations
* p < 0.1, ** p < 0.05, *** p < 0.01
The following are the same

### 4.3 Endogeneity Test

The baseline regression verifies the boosting effect of ESG on Inv, but this result may be due to the reverse causality brought about by the ability and willingness of the more efficient investment firms to improve their ESG performance. Considering that lagged-period ESG is not susceptible to the reverse effects of current period Inv, this paper re-runs the regression test using ESG with one, two, and three lags (L.ESG, L2.ESG, L3.ESG) as explanatory variables respectively to mitigate the reverse causality. According to the test findings of Table 4, Inv and ESG are considerably negative at the 1% level for each lag, i.e., corporate ESG performance has a longer-term and sustained impact on enhancing corporate investment efficiency. After considering the endogeneity problem, the results of this research remain valid.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L.ESG</td>
<td>-0.003***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L2.ESG</td>
<td></td>
<td>-0.004***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td>L3.ESG</td>
<td></td>
<td></td>
<td>-0.003***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.001)</td>
</tr>
</tbody>
</table>
4.4 Robustness Test

This paper refers to the method of Gao et al. (2021)\(^2\) to reconstruct the explanatory variable ESG2, with ESG2=1 when the ratings are C–CCC, ESG2=2 when the ratings are B–BBB and ESG2=3 when the ratings are A–AAA. After replacing the explanatory variables, columns (1) and (2) of Table 5 shows the results, the significant levels of ESG2 and Inv remain at 1%, which is the same with previous findings, and the test results of hypothesis H1 can be considered robust.

Second, because the 2020 epidemic may have an unanticipated effect on corporate ESG performance, this article omits the sample data from 2020 and re-estimates the link between ESG and Inv using the sample data from 2012 to 2019. Columns (3) and (4) of Table 5 display the regression results, where ESG and Inv remain at the 1% level and the regression results do not change substantially.

<table>
<thead>
<tr>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ESG2 -0.009</td>
<td>-0.007 ***</td>
<td>-0.004 ***</td>
<td>-0.004 ***</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.002)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>ESG -0.004</td>
<td>-0.004 ***</td>
<td>-0.004 ***</td>
<td>-0.004 ***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td>Constant 0.043</td>
<td>0.088 ***</td>
<td>0.047 ***</td>
<td>0.082 ***</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.028)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>Controls NO</td>
<td>Yes NO</td>
<td>Yes NO</td>
<td>Yes NO</td>
</tr>
<tr>
<td>Industry Yes</td>
<td>Yes Yes</td>
<td>Yes Yes</td>
<td></td>
</tr>
<tr>
<td>Year 9684</td>
<td>9684 8608</td>
<td>8608 8608</td>
<td></td>
</tr>
<tr>
<td>R² 0.054</td>
<td>0.092 0.046</td>
<td>0.055 0.046</td>
<td>0.093 0.084</td>
</tr>
<tr>
<td>Adj.R² 0.047</td>
<td>0.084 0.046</td>
<td>0.046 0.046</td>
<td>0.084 0.084</td>
</tr>
</tbody>
</table>

5 MODERATING EFFECT TEST

5.1 Moderating Effect of Soe

The results of column (1) of Table 6 show that the ESG × Soe passes the significance test at the 5% level. The moderating impact of Soe on the effect of ESG on Inv is convincingly proven, and hypothesis H2 is tested. Further analysis reveals that the extent to which ESG performance enhances investment efficiency decreases in a state-owned enterprise. The reason for this is that
state-owned corporations bear more responsibilities than non-state-owned corporations for economic development, environmental protection, and promoting social harmony, they may make responsible investments that are not very helpful to the long-term development, while they can achieve higher ESG ratings, it is challenging to increase the enterprise's investment efficiency.

5.2 Moderating Effect of Media

The results in column (2) of Table 6 show that the ESG×Media passes the significance test at the 10% level. It is sufficient to show that Media has a negative moderating effect, and hypothesis H3 is verified. Further research demonstrates that for corporations with high media attention, it has negative effects on the contrary. The reason for this is that the media may focus on negative information about ESG underperforming firms, passive ESG practices may disrupt firms' original investment plans, resulting in more inefficient investments.

<table>
<thead>
<tr>
<th>Table 6: Moderating Effect Test of external pressure.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
</tr>
<tr>
<td>ESG</td>
</tr>
<tr>
<td>(0.001)</td>
</tr>
<tr>
<td>Soe</td>
</tr>
<tr>
<td>(0.006)</td>
</tr>
<tr>
<td>ESG×Soe</td>
</tr>
<tr>
<td>(0.001)</td>
</tr>
<tr>
<td>Media</td>
</tr>
<tr>
<td>(0.002)</td>
</tr>
<tr>
<td>ESG×Media</td>
</tr>
<tr>
<td>(0.000)</td>
</tr>
<tr>
<td>Constant</td>
</tr>
<tr>
<td>(0.023)</td>
</tr>
<tr>
<td>Controls</td>
</tr>
<tr>
<td>Industry</td>
</tr>
<tr>
<td>Year</td>
</tr>
<tr>
<td>N</td>
</tr>
<tr>
<td>R²</td>
</tr>
<tr>
<td>Adj.R²</td>
</tr>
</tbody>
</table>

6 CONCLUSION AND RECOMMENDATION

This paper constructs a multiple linear regression model to test the linear relationship of corporate ESG performance and corporate investment efficiency while examining the moderating effect of external pressure that exists between the two. The empirical findings demonstrate that ESG performance significantly improves the effectiveness of corporate investment. The moderating mechanism reveals that external pressure has a significant moderating effect between ESG and Inv, indicating that ESG practices driven by external pressure reduce corporate investment efficiency due to the specificity of the Chinese capital market, resulting in counterproductive effects.
The following suggestions are made in light of the aforesaid results: Firstly, corporations should pay attention to the role of ESG performance in improving investment efficiency, and promote the in-depth application of ESG investment concepts in enterprise investment decisions; Secondly, the government should formulate pertinent policies and regulations to guide corporations in optimizing their investment structures, and should provide them specific incentives or appropriate penalties in terms of tax policies, loan projects, and government bidding projects according to their ESG performance; Thirdly, state-owned corporations should fully consider the degree of emphasis of different indicators in ESG ratings and take multiple measures to improve their investment efficiency. Fourthly, the media should maintain benign and close attention to corporations with insufficient ESG performance, properly play the role of public opinion supervision, maintain the independence of the news industry, and at the same time actively report positive information on the ESG performance.

REFERENCES


International Oil Price Uncertainty, Operating Net Cash Flow and Corporate Investment Analysis Based on Stata and Python

Yi Wang1,a*, Ye Miao1,b and Xu Liang1,c

*2290847993@qq.coma, 1720785905@qq.comb, ijoan59@163.comc

School of Economics and Management, Beijing Jiaotong University, Beijing 100044, China1

Abstract: Under the influence of supply and demand, politics and other factors, international oil prices have been in sharp fluctuations. As a large crude oil consumption country, oil prices frequent fluctuations will affect China's economic growth. It is important to study the microeconomic effects of oil price fluctuations. Based on this, this paper downloads the company's data from CASMA, and uses Python software to crawl missing data from the annual financial report. Then this paper takes the panel data of all A-share listed companies from the first quarter of 2014 to the fourth quarter of 2019 as samples and uses Stata software to build OLS regression model to study the impact of international oil price uncertainty on corporate investment. The empirical results show that: (1) International oil price uncertainty has a significant inhibitory effect on the investment; (2) The negative effect between two is partly caused by the decrease of operating net cash flow; (3) International oil price uncertainty has a more significant inhibitory effect on small-scale corporate investment.

Keywords: International Oil Price Uncertainty, Corporate Investment, Operating Net Cash Flow, Stata, Python.

1 INTRODUCTION

Economic growth and the continuous expansion of enterprises continue to drive the demand for oil. From 2019 to 2021, China's annual crude oil imports exceeded 500 million tons, which is the world's largest oil importing country. The high dependence on crude oil imports makes China’s economy very sensitive to oil price changes. In addition, with the reform of the pricing mechanism of refined oil products, the relationship between domestic crude oil prices and international crude oil prices is increasingly close. However, in recent years, international oil price has fluctuated frequently under the influence of supply and demand, macroeconomic policies and politics, which has exerted a great influence on our economic growth and development. In recent years, China has vigorously deployed new energy industries to replace traditional energy consumption. However, now China's energy system has problems such as the bottleneck of new energy development and overcapacity of traditional energy. A reliable solution to these problems is to rely on the information technology industry. The risks brought
by oil price fluctuations will affect the smooth operation of the stock market. At this stage, new energy and information technology enterprises will be greatly affected by oil price fluctuations.

As one of the important factors of production in enterprises, the wild fluctuation of crude oil price has a great impact on the development of Chinese macro and micro economy. Crude oil is the direct or indirect input cost of most enterprises. The uncertainty of its price makes the investment decision of enterprises more difficult. Besides, corporate investment is also one of the important channels for crude oil price to be transmitted to the macro economy. Under the complicated economic situation, it is of great practical significance to study the impact of international oil price on corporate investment and its transmission mechanism, which is helpful to improve the economic benefits of enterprises and promote the healthy development of our real economy. Financing constraint is an important factor affecting enterprise investment as well, which is largely determined by enterprise scale. Therefore, the influence of enterprise size on the relationship between the two is also the focus of this paper.

2 LITERATURE REVIEW AND HYPOTHESIS DEVELOPMENT

2.1 Literature Review

Scholars have explored some research on the impact of oil price uncertainty on micro-firm behaviour, focusing on corporate performance, cash holdings, and investment. Oil price uncertainty leads to a substantial increase in the uncertainty of enterprises’ production and operation, which leads to a decline in corporate performance [11]. When oil price uncertainty ascends, companies tend to increase their cash holdings for precautionary reasons. Cash holdings can reduce external financing costs and hedge future investment risks, which is of great significance for enterprises to deal with oil price uncertainties [15]. Some studies have also found that there is an inverted U relationship between oil price uncertainty and corporate cash holdings [17]. Companies tend to view investments as a waiting option, and uncertainty increases the value of their investments in waiting options. Henriques et.al (2011) [6] found a positive U-shaped relationship between oil price uncertainty and corporate investment. Some scholars have found a negative correlation between the two. Han (2016) [5] found a significant negative correlation between the two in manufacturing firms. Yang (2018) [18] found a significant negative correlation between the two in the oil industry.

Corporate investment is affected by a variety of factors, not only by the internal conditions of the company, such as capital structure and profitability, but also by the impact of the external environment, such as macroeconomic impact and external environmental uncertainty. The net cash flow generated from operating activities is an important pillar of enterprise development. The more cash flow an enterprise generates in its business activities, the stronger its willingness to invest, and this influence is more significant in small and medium-sized enterprises [9]. Most enterprises use debt for investment. Enterprises with excessive debt face high operating risks and debt repayment pressure, and are also subject to large financing constraints, which results in a very low willingness to invest [7]. From external conditions, GDP growth reflects a country’s economic progress and consumer spending power, the faster GDP grows, the more enterprises are inclined to invest [4]. External environment uncertainty leads to a substantial increase in the
business and financial risks of enterprises, and the cash flow expected to be generated by investment fluctuates greatly, which leads to the decision of enterprises to delay investment. Most studies find that uncertainty has a negative impact on enterprise investment.

To sum up, studies on the relationship between oil price uncertainty and investment mainly focus on the direct impact. A few studies analyse the mechanism of the impact of international oil price uncertainty on investment. In addition, the relationship between oil price uncertainty and investment is also affected by many other factors, such as scale, but relevant studies are few and mainly focused on a certain industry, such as manufacturing and petroleum industry.

2.2 Theoretical Analysis and Hypothesis

Crude oil is one of the basic inputs for the production of most goods and services. Although some companies may not directly consume or produce crude oil, in most cases crude oil can be regarded as an indirect cost for the company. The sharp fluctuation of oil price leads to greater uncertainty for enterprises. Based on the real option theory, when an enterprise is faced with a high degree of uncertainty, the cash flow generated by investment in the future is highly volatile, and the value of the enterprise's investment waiting option will rise significantly (Bernanke, 1983). From the perspective of financing, on the one hand, oil price uncertainty increases the enterprises’ risk and information asymmetry, so banks and other credit institutions will increase the lending rate [3]. On the other hand, oil price uncertainty leads to increased volatility in corporate stock prices and the necessary yield demanded by investors will also rise. Together, corporate investment will fall.

Hypothesis 1: International oil price uncertainty has a significant inhibitory effect on the investment.

Oil price fluctuations can lead to substantial declines in consumer spending expectations, especially for consumer durables. The decrease of consumer demand will lead to the decline of enterprise product sales and the deterioration of corporate performance [10-11]. Besides, in order to cope with oil price uncertainty, enterprises will take a series of measures to reduce risks, which will undoubtedly increase the capital expenditure of enterprises in the current period. Under the joint action of the two, corporate operating net cash flow will decline. According to the neoclassical investment theory, when making investment, enterprises usually consider whether the internal capital is sufficient first. The more internal capital, the stronger the investment intention. Cash flow in business activities is the basis of sustainable operation of enterprises and an important source of enterprise funds, which can reflect the financial situation of enterprises to a certain extent. The more abundant the cash flow, the less external financing constraints it faces, the stronger the investment will be [8].

Hypothesis 2: The negative effect of international oil price uncertainty on corporate investment is partly caused by the decrease of operating net cash flow.

External financing constraints are also important factors affecting corporate investment. The capital strength of small-scale enterprise is weak, and the ability to resist risks is also low. Banks, investors and others are reluctant to provide more capital to small companies for safety reasons, or demand a big increase in the necessary return rate. Therefore, in the face of external uncertainties, small companies are more inclined to retain funds and reduce investment for security reasons. In contrast, large-scale enterprises have very strong financial strength, more
resources, and more financing channels. In addition, large-scale enterprises have the ability to improve the skill level of their employees and thus improve the efficiency of resource utilization. It is also possible to hedge against adverse cost changes by buying futures or other energy derivatives hedging instruments. Although the rise in international oil prices will lead to a decrease in corporate profits and a small decrease in investment to some extent, the extent is weaker than that of small-scale enterprises.

Hypothesis 3: International oil price uncertainty has a more significant inhibitory effect on small-scale corporate investment.

3 METHODS AND MATERIALS

3.1 Data Processing

This paper selects Chinese A-share listed companies from the first quarter of 2014 to the fourth quarter of 2019 as samples for research. The sample cut-off time is December 31, 2019 for the following reasons: the COVID-19 pandemic has had a huge impact on the production and operation of enterprises from 2020, and the investment of most enterprises has shown a downward trend since 2020. The impact of the COVID-19 pandemic on corporate investment has seriously interfered with the impact of oil price fluctuations on the investment.

The company data studied in this paper is mainly from CSMAR and oil price data is from the U.S. Energy Information Administration. Since all the data in this paper are quarterly data, the data directly obtained from the website is seriously missing. So we also use Python software to crawl the missing data from the company's annual report.

Then we use Stata software to integrate data and build OLS regress model perform multi-regression analysis. The procedures for processing data with computer software are as follows. First, import the acquired data into Stata software and convert it from “xlsx” format to “dta” format; Second, use the “merge” and “append” commands to integrate data from different sources to form a total “dta” file; Third, use “tsset” “gen” and other commands to process data and generate proxy variables required by this paper; Fourth, use “drop” commands to exclude the data with financial industry, with total assets less than 0 and listed for less than one year. Fifth, use the “winsorize” command to shrink the 1% and 99% of company level data. After processing, we obtained a total of 58,280 company quarters of unbalanced panel data. Finally, use the “sum” and “reg” commands to perform descriptive statistics and regression analysis on panel data.

3.2 Variable Definitions

Following Wang (2017), we use “Funds paid for the purchase and construction of fixed assets, intangible assets and other long-term assets” as the proxy variable of corporate investment. Following Phan et al. (2019)[9], calculating the quarterly standard deviation based on the logarithmic return of WTI daily crude oil price \( Z_i = \ln(p_i / p_{i-1}) \) and the actual trading days of each quarter. And using the mean standard deviation with a lag of 1 to 4 periods as the proxy variable of oil price uncertainty.
\[ \delta_i = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (Z_i - E(Z_i))^2} \]  

(1)

\[ SDOIL_i = (\delta_{i,1} + \delta_{i,2} + \delta_{i,3} + \delta_{i,4}) / 4 \]  

(2)

Referring to the existing literature, this paper controls variables such as corporate leverage, cash flow, growth capacity to consider the firm related heterogeneity that affects corporate investment.

**Table 1: Variables description**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent variables</td>
<td>Invest</td>
<td>&quot;Funds paid for the purchase and construction of fixed assets, intangible assets and other long-term assets&quot;/ total assets at the beginning of the quarter</td>
</tr>
<tr>
<td>Intermediate variable</td>
<td>CF</td>
<td>Operating net cash flow/ total assets at the beginning of quarter</td>
</tr>
<tr>
<td>Independent variables</td>
<td>SDOIL</td>
<td>Reference model (1) and (2)</td>
</tr>
<tr>
<td></td>
<td>TQ</td>
<td>Tobin Q</td>
</tr>
<tr>
<td></td>
<td>Lev</td>
<td>total liabilities/ total assets at the beginning of the quarter</td>
</tr>
<tr>
<td></td>
<td>Growth</td>
<td>Growth rate of revenue</td>
</tr>
<tr>
<td></td>
<td>Size</td>
<td>The natural log of total assets</td>
</tr>
<tr>
<td></td>
<td>TOP5</td>
<td>Shareholding ratio of the top five shareholders</td>
</tr>
<tr>
<td></td>
<td>Management</td>
<td>overhead/ Prior quarter sales revenue</td>
</tr>
<tr>
<td></td>
<td>Cash</td>
<td>Cash and cash equivalents/ Total assets</td>
</tr>
<tr>
<td></td>
<td>GDP</td>
<td>GDP growth rate</td>
</tr>
<tr>
<td></td>
<td>Ndts</td>
<td>overhead and expense of sales /Prior quarter sales revenue</td>
</tr>
<tr>
<td></td>
<td>SIZE</td>
<td>Dummy Variables: equal 1 if firm size is larger than the median</td>
</tr>
</tbody>
</table>

### 3.3 Model Design

Following Phan et al. (2019)[9], this paper use Stata software to build OLS multiple regression model (3) to test hypothesis 1. Except for international oil price uncertainty, the model (3) also controls nine additional variables, including variables at the enterprise level, such as Q, Lev etc., as well as macro-level influences, such as GDP.

\[ \text{Invest}_{it} = \beta_0 + \beta_1 \text{SDOIL}_{it} + \beta_2 Q_{it} + \beta_3 \text{Cash}_{it} + \beta_4 \text{Growth}_{it} + \beta_5 \text{Lev}_{it} + \beta_6 \text{Size}_{it} + \beta_7 \text{TOP5}_{it} + \beta_8 \text{Management}_{it} + \beta_9 \text{GDP}_{it} + \beta_{10} \text{Ndts}_{it} + \sum \text{QUARTER} + \sum \text{YEAR} + \varepsilon_{it} \]  

(3)
where \( i \) is firm and \( t \) is time. \( \beta_0 \) control company fixed effects. Model (3) also includes annual and quarterly dummy variables. In this model, \( \beta \) reflects the influence of international oil price uncertainty on corporate investment, \( \beta_i < 0 \) verify the hypothesis 1 is true.

Referring to the research design of Xue (2020), “operating net cash flow” is taken as the proxy variable of \( CF \). Following Wen (2014), we establish models (4) and (5) and form a recursive model with model (3) to test hypothesis 2.

\[
CF_{it} = \alpha_0 + \alpha_i SDOIL_{it} + \alpha_i Q_{it} + \alpha_i Cash_{it} + \alpha_i \text{Growth}_{it} + \alpha_i \text{Lev}_{it} + \alpha_i \text{Size}_{it} + \sum QUARTER + \sum \text{YEAR} + \epsilon_{it} \tag{4}
\]

\[
Invest_{it} = \gamma_0 + \gamma_i SDOIL_{it} + \gamma_i CF_{it} + \sum Controls_i + \sum QUARTER + \sum \text{YEAR} + \epsilon_{it} \tag{5}
\]

where \( i \) is firm and \( t \) is time. Both model (4) and Model (5) control firm fixed effect and time effect. Based on the expectation of hypothesis 2, \( \alpha_1 \) in the model (4) is negative, \( \gamma_2 \) is positive, and \( \gamma_1 \) is negative in the model (5).

This paper divides the enterprise scale according to the total assets and set the dummy variable of \( SIZE \). Enterprises whose total assets are higher than the average level are classified as large-scale enterprises and \( SIZE=1 \). On the contrary, it is a small-scale enterprise and \( SIZE=0 \). Based on the research design of Phan (2019)\(^{[9]} \) and Aktham (2020)\(^{[1]} \), we establish model(6) to test hypothesis 3.

\[
Invest_{it} = \phi_0 + \phi_i SDOIL_{it} + \phi_i SIZE_{it} + \phi_i SDOIL_{it} \times SIZE_{it-1} + \sum Controls_i + \sum QUARTER + \sum \text{YEAR} + \epsilon_{it} \tag{6}
\]

If \( \phi_1 \) is negative and \( \phi_3 \) is positive, it can be verified that firm size has a positive moderating effect.

4 REGRESSION ANALYSIS

4.1 Descriptive Statistics

Table 2 presents the results of descriptive statistics after tailoring (1% before and after) each variable. \( Invest \) is a relative value that indicates how much capital expenditures account for total assets, and it fluctuates widely. The mean value is 0.013, indicating most firms invest about 1% of total assets, which is relatively low overall. \( SDOIL \) also varies considerably between quarters, with some quarters having very high oil price volatility during the sample period.
Table 2: Descriptive statistic

<table>
<thead>
<tr>
<th>variable</th>
<th>N</th>
<th>Mean</th>
<th>Median</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Invest1</td>
<td>58280</td>
<td>0.013</td>
<td>0.008</td>
<td>0</td>
<td>0.085</td>
</tr>
<tr>
<td>SDOIL</td>
<td>58280</td>
<td>0.082</td>
<td>0.080</td>
<td>0.043</td>
<td>0.129</td>
</tr>
<tr>
<td>CF</td>
<td>58280</td>
<td>0.013</td>
<td>0.012</td>
<td>-0.102</td>
<td>0.149</td>
</tr>
<tr>
<td>Q</td>
<td>58280</td>
<td>2.139</td>
<td>1.733</td>
<td>0.892</td>
<td>8.353</td>
</tr>
<tr>
<td>Cash</td>
<td>58280</td>
<td>0.143</td>
<td>0.112</td>
<td>0.010</td>
<td>0.555</td>
</tr>
<tr>
<td>Growth</td>
<td>58280</td>
<td>0.138</td>
<td>0.036</td>
<td>-0.799</td>
<td>3.698</td>
</tr>
<tr>
<td>Lev</td>
<td>58280</td>
<td>0.414</td>
<td>0.401</td>
<td>0.053</td>
<td>0.888</td>
</tr>
<tr>
<td>Size</td>
<td>58280</td>
<td>22.20</td>
<td>22.03</td>
<td>19.95</td>
<td>26.17</td>
</tr>
<tr>
<td>TOP5</td>
<td>58280</td>
<td>54.24</td>
<td>54.51</td>
<td>21.19</td>
<td>88.39</td>
</tr>
<tr>
<td>Management</td>
<td>58280</td>
<td>0.101</td>
<td>0.078</td>
<td>-0.037</td>
<td>0.604</td>
</tr>
<tr>
<td>GDP</td>
<td>58280</td>
<td>0.027</td>
<td>0.090</td>
<td>-0.164</td>
<td>0.117</td>
</tr>
<tr>
<td>Nds</td>
<td>58280</td>
<td>0.180</td>
<td>0.137</td>
<td>0.004</td>
<td>0.853</td>
</tr>
</tbody>
</table>

4.2 Regression Result Analysis

The first column of Table III shows the coefficient $\beta_1$ between SDOIL and Invest1 is -0.014, and is significant at 1% level, indicating oil price uncertainty has a significant inhibitory effect on corporate investment, and verifying Hypothesis 1 is true. The coefficients between $Q$, Cash, Growth and Invest1 are significantly positive, indicating enterprises with stronger profitability, more cash and higher operating income are more willing to invest. The coefficient between Lev and Invest1 is significantly negative, indicating enterprises are reluctant to invest when debt burden is high, and will reduce the scale of investment.

The second and third columns of Table III show the coefficient $\alpha_1$ between SDOIL and CF is significantly negative at the 5% level. The coefficient $\gamma_2$ between CF and Invest1 is significantly positive, indicating oil price uncertainty reduces investment through the channel of reducing corporate net cash flow from operations. Hypothesis 2 is supported. Specifically, every increase of one unit of oil price uncertainty decreases operating net cash flow by 0.0266 units on average, and every increase of one unit of operating net cash flow increases corporate investment by another 0.0229 units on average. The coefficient symbols of control variables and investment in the model (5) are basically the same as those in the model (3).

Table 3: Empirical results of hypothesis 1 and 2

<table>
<thead>
<tr>
<th></th>
<th>Invest1</th>
<th>CF</th>
<th>Invest1</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDOIL</td>
<td>-0.0140***</td>
<td>-0.0266</td>
<td>-0.0137***</td>
</tr>
<tr>
<td></td>
<td>(-3.00)</td>
<td>(-1.89)</td>
<td>(-2.96)</td>
</tr>
<tr>
<td>CF</td>
<td></td>
<td>0.0229**</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(15.48)</td>
<td></td>
</tr>
<tr>
<td>Q</td>
<td>0.001***</td>
<td>0.001***</td>
<td>0.001***</td>
</tr>
<tr>
<td></td>
<td>(8.64)</td>
<td>(4.03)</td>
<td>(8.34)</td>
</tr>
</tbody>
</table>
Table IV shows the coefficient of SDOIL and Invest1 is significantly negative, while the coefficient of SDOIL × SIZE is significantly positive, indicating firm size alleviates the negative effect of oil price uncertainty on investment. Hypothesis 3 is supported. The coefficient of another control variables is the same as hypothesis 1.

**Table 4**: Empirical results of hypothesis 3

<table>
<thead>
<tr>
<th></th>
<th>Invest1</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDOIL</td>
<td>-0.018*** (-3.59)</td>
</tr>
<tr>
<td>SDOIL × SIZE</td>
<td>0.008** (2.06)</td>
</tr>
<tr>
<td>Controls</td>
<td>Yes</td>
</tr>
<tr>
<td>Firm/Time FE</td>
<td>Yes</td>
</tr>
<tr>
<td>r2</td>
<td>0.055</td>
</tr>
<tr>
<td>N</td>
<td>58,280</td>
</tr>
</tbody>
</table>

### 4.3 Robustness Tests

This paper uses the alternative measures of oil price uncertainty to test robustness. Following Sadorsky (2006) and Wang (2017), for the uncertainty of international oil prices, this paper uses

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Firm/Time FE Yes</td>
</tr>
<tr>
<td></td>
<td>r2        0.055</td>
</tr>
<tr>
<td></td>
<td>N         58,280</td>
</tr>
</tbody>
</table>
GARCH(1,1) to estimate the volatility of the logarithmic return of WTI daily crude oil spot price, and then takes the arithmetic mean value of the quarter with a lag of 1 to 4 periods as the proxy variable of international oil price uncertainty (GARCHOIL).

\[ e_t = Z_t + 0.0256Z_{t-1} - 0.0146Z_{t-2} + 0.0368Z_{t-3} - 0.0091Z_{t-4} - 0.0187Z_{t-5} - 0.0188 \]  

(7)

\[ h_t = 0.0351 - 0.0630e_{t-1}^2 - 0.9317h_{t-1}^E \]  

(8)

\[ \text{GARCHOIL}_{it} = (h_{t-1}^a + h_{t-2}^a + h_{t-3}^a + h_{t-4}^a)/4 \]  

(9)

Specifically, the measurement period of GARCH (1,1) model is from January 1, 2013 to December 31, 2019. When choosing the order of auto-regressive lag, this paper found that the data with lag period 5 passed the most criteria, so the paper adopted the model with lag period 5. The specific GARCH model is shown in equations (7) and (8). \( h_t \) represents the daily yield fluctuation amplitude predicted by GARCH (1,1) model, and \( e_t \) is the residual of model (8). \( Z_t \) is the logarithmic return of WTI daily crude oil price \((1 \ln(p_t / p_{t-1}))\). The quarterly oil price uncertainty \( h_t^E \) is the quarterly average of the daily yield fluctuation range \( h_t \).

**Table 5: Robustness test results of hypothesis 1 and 2**

<table>
<thead>
<tr>
<th></th>
<th>Invest1</th>
<th>CF</th>
<th>Invest1</th>
</tr>
</thead>
<tbody>
<tr>
<td>GARCHOIL</td>
<td>-1.161***</td>
<td>-3.015**</td>
<td>-1.131***</td>
</tr>
<tr>
<td></td>
<td>(-2.74)</td>
<td>(-2.41)</td>
<td>(-2.67)</td>
</tr>
<tr>
<td>CF</td>
<td></td>
<td></td>
<td>0.023***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(15.48)</td>
</tr>
<tr>
<td>Controls</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>r2</td>
<td>0.055</td>
<td>0.176</td>
<td>0.059</td>
</tr>
<tr>
<td>N</td>
<td>58,280</td>
<td>58,280</td>
<td>58,280</td>
</tr>
</tbody>
</table>

The coefficient of \( \text{GARCHOIL} \) and \( \text{Invest1} \) is significantly negative, while the coefficient of \( \text{GARCHOIL} \) and \( \text{CF} \) is significantly negative and the coefficient of \( \text{CF} \) and \( \text{Invest1} \) is significantly positive, which is consistent with the above regression results.

**Table 6: Robustness test results of hypothesis 3**

<table>
<thead>
<tr>
<th></th>
<th>Invest1</th>
</tr>
</thead>
<tbody>
<tr>
<td>GARCHOIL</td>
<td>-1.580***</td>
</tr>
<tr>
<td></td>
<td>(-3.34)</td>
</tr>
<tr>
<td>GARCHOIL*SIZE</td>
<td>0.777**</td>
</tr>
<tr>
<td></td>
<td>(2.00)</td>
</tr>
<tr>
<td>Controls</td>
<td>Yes</td>
</tr>
<tr>
<td>r2</td>
<td>0.055</td>
</tr>
<tr>
<td>N</td>
<td>58,280</td>
</tr>
</tbody>
</table>
The coefficient of $GARCH_{OIL}$ and $Invest1$ is significantly negative and the coefficient of $GARCH_{OIL} \times SIZE$ and $Invest1$ is significantly positive. This result supports hypothesis 3, which can fully verify size alleviates the inhibitory effect between two.

5 CONCLUSIONS

This paper uses the complete data from the first quarter of 2014 to the fourth quarter of 2019 that are downloaded from CASAM and EIA and that obtained from Python crawlers, then uses Stata software to build an OLS regression model to empirically analyse the relationship between international oil price uncertainty and corporate investment. The results show that international oil price uncertainty has a significant inhibitory effect on corporate investment. The uncertainty of the international oil price will cause the decrease of operating net cash flow, then inhibit the investment of enterprises. The size of enterprises mitigated the negative effect of the uncertainty of international oil price on investment.

Based on the above conclusions, this paper has important practical implications. First of all, the government should continuously improve the stability of price transmission and the flexibility of price adjustment, so as to reduce the severe impact of oil price uncertainty on enterprises and the harm to the real economy. Secondly, the establishment of domestic oil security early warning mechanism, continue to improve the domestic oil price volatility hedging system, maximize the futures market hedging and risk management functions. Finally, in the critical period of national transformation, enterprises should constantly improve their own risk management system to effectively cope with external shocks and maintain steady development of enterprises, especially new energy and information technology enterprises.

REFERENCES

Design and Implementation of Entrepreneur Evaluation and Analysis System Based on AHP Method

Na Guo1,a, Xiaoxian Gong1,b

guonasdoar@163.com,a, 455628892@qq.comb

Dongfeng College, Shandong University of Finance and Economics, Tai'an City, Shandong Province, 271000, China1

Abstract: The evaluation of entrepreneurs has a great influence on the brand awareness and the attractiveness of investors, which in turn can affect the performance of an enterprise. Therefore, this paper uses java web technology to develop an entrepreneur evaluation and analysis system, systematically uses AHP method and designs related algorithms to analyze the influencing factors of entrepreneur evaluation, and systematically judges entrepreneur evaluation in a scientific way according to the judgment matrix and the results of

Keywords: AHP, Java web, Entrepreneur Evaluation, Application System Design.

1 INTRODUCTION

With the continuous development of market economy and the continuous improvement of China's modern enterprise system, more and more private entrepreneurs emerge in an endless stream, contributing to the development of China's economic market, and the talents of entrepreneurs have strategic marketing for China's current economy and society. Under the background of the Internet technology era, with the rapid development of the Internet, the evaluation of entrepreneurs is becoming more and more networked, and the evaluations of entrepreneurs spread rapidly on the Internet, which will have an important impact on brand reputation. Therefore, it is of great significance for the development of China's market economy to study the evaluation of Chinese entrepreneurs and deepen the reform of enterprises. In China's existing entrepreneur evaluation, the method is traditional, the index system is not comprehensive, and there are some problems of pertinence and intersection. Therefore, we need to study a more scientific evaluation system for entrepreneurs. According to the above content analysis, the author of this paper thinks that the Internet technology and AHP method should be used to study the entrepreneur evaluation and analysis system. The indicators come from the network and are displayed in the way of application system, which can reflect the evaluation of entrepreneurs more intuitively, reveal the evaluation status of entrepreneurs, find out their possible strengths and weaknesses, and provide empirical basis for consumers and investors. [1]
2 KEY TECHNOLOGY

2.1 Web

Web development technology mainly refers to web page development technology, which is a front-end web page development technology. The technical front-end is mainly developed by using the technical framework of HTML+CSS+JavaScript. JavaScript is an explanatory scripting language, and JavaScript is the most widely used front-end development technology, so most browsers have built-in JavaScript interpreters. The common tools of front-end development are development frameworks such as VUE.js, React and bootstrap to improve the efficiency of front-end page development. Especially under the technical premise of using HTML5 and CSS3, the page interaction and visual aesthetic effect can be greatly enhanced.

Compared with HTML, HTML5 has the advantage of adding a larger number of tag libraries and optimizing the functional design of form processing, which makes HTML5 technology more adaptable in the current era. In addition, HTML5 has the advantages of reducing the dependence of external plug-ins and supporting localized offline storage mode. Similarly, CSS3 is also an upgraded version of CSS. Compared with CSS, CSS3 technology has more module effects, such as language, list, background border and text effects. The advantage of increasing module special effects is that it can greatly reduce the number of HTTP requests, and at the same time reduce the redundancy of tags, thus enhancing the user experience. Web front-end development technology needs to combine ajax technology to realize the complete interaction of pages, and at the same time construct a bridge between the client and the server for asynchronous response. [2]

2.2 Java

Java is an interpretive development language based on object-oriented thinking. Java features include distributed, efficient and robust, cross-platform and open source. At the same time, Java also supports network programming and multi-thread programming, so Java has gradually become the most mainstream development language. C++ is the most widely used programming language in the previous generation. Compared with C++, Java abandons complex functions such as pointer and memory management, thus simplifying the development process and further improving the development efficiency. But Java is still a class-based development language, which has three characteristics: inheritance, encapsulation and polymorphism. Java language can inherit the old code to improve the reusability of the code; Encapsulation features can isolate the code into small segments, making it easier to use; Polymorphism can make the code more flexible in use. And because the program developed by Java language is dynamic, the application system developed by Java language can be upgraded freely in a distributed environment, and it is not limited by the application of the source program.

The Java application development process is shown in Figure 1. First, developers write the code of the source file, and then compile the source file with a compiler. In this process, bytecodes will be generated, and the bytecodes can be executed by an interpreter to complete the application development process. Based on the above analysis, the research and development of entrepreneur evaluation and analysis system is carried out in this paper based on JavaWeb. [3]
2.3 AHP Method

Analytic hierarchy process (AHP) comes from professor Thomas L. Saaty in the United States. This analysis method is applied to the analysis of decision-making problems with multiple evaluation criteria and uncertain conditions. Decision makers can use AHP method to determine the priority of evaluation criteria and quantify decision variables according to their importance and decision variables, so as to calculate the best decision-making mode scientifically. The characteristic of AHP is that it can use less and quantitative information to mathematically calculate the thinking process of decision-making. In this process, the decision-makers need to deeply analyze the essential influencing factors and their internal relations of the problems in the decision book. AHP analytic hierarchy process can simplify the problems with multi-objectives, multi-criteria or no structural characteristics. Therefore, this paper uses AHP analytic hierarchy process to analyze entrepreneurs' comprehensive evaluation. [4]

2.4 Development Environment

The development language is JAVA 8, the development tool is IntelliJ IDEA, and the server uses Apache Tomcat8.0. The front-end development language of the system is HTML+CSS+JavaScript, the development tool is Vue.js combined with bootstrap, and ajax asynchronous request technology is used to complete the data interaction between the front and back ends. The page developed by ajax technology can refresh locally, which can improve the loading speed of the page and enhance the user experience. The back-end development framework of the system chooses the SSM framework structure of spring+springmvc+mybatis, which is the current mainstream. The choice of system architecture and key implementation technologies is very important. According to the analysis of the relevant technologies in the current era, it is considered that it is technically feasible to build an entrepreneur evaluation and analysis system based on the above technologies.

3 DEVELOPMENT PROCESS

The application system of entrepreneur evaluation and analysis based on AHP method is designed based on Javaweb technology, and the technical architecture of the system is shown in Figure 2. The system functions are designed in a hierarchical way. The system architecture
The system is developed with the front-end separation mode, and the front-end framework adopts VUE, which is responsible for the view layer in MVC. Yii is responsible for the control layer and template layer of MVC pattern. The front-end framework adopts VUE, the front-end development language uses HTML+JS+CSS, and the static files of picture documents are saved. The display layer is rendered by browser and interacted with ajax technology. The data layer is responsible for the connection between the system and the database, including data backup, database reading and writing, information storage. Mybatias is the persistence layer framework, which calls the DAO interface to POJO JAVA objects. Spring is a framework for coupling business layer and other layers, and it is a lightweight java development framework. The configuration of springboot version can be better simplified. The MVC module of this system is built by Springmvc framework, which divides the model, filter and controller, making it easier to customize the system. [5]
4 FUNCTIONAL IMPLEMENTATION

After the user logs in through the account password, the entrepreneur's rating can be obtained by inputting the entrepreneur information specified by the system. The entrepreneur evaluation system needs high authority to access, so the function of judging the legality of user identity should be maintained. After the user inputs the corresponding account number and password data, the server receives the login request and calls the relevant interface required by the login function to verify the information data input by the user. In this paper, the class of login function is described, and the class diagram of login function is shown in Figure 3. You can see that the login function is mainly set in five categories. UserInfo is the main class, which is mainly responsible for the management of all kinds of basic information of users, and holds the attribute information of account and password required by users to log in. The LoginController class is used to handle the user's requests to the system, while the LoginService class is mainly used to handle the user's requirements in business logic. The LoginDao class is the related attribute that is responsible for the interaction between the database server and the system. Finally, the CheckFilter class, whose function is to verify the permissions of all kinds of requests issued by users and intercept erroneous requests in time.

![Figure 3: The accession functional class diagram](image)

The enterprise price evaluation method of this system is AHP, which uses JAVA language to realize logic. In this part, the author will elaborate the concrete steps of AHP method. First of all, it is necessary to analyze the relationship between the factors in the entrepreneur evaluation system and establish the system. This paper adopts hierarchical structure. The hierarchical structure of this paper consists of three layers: the first layer is the target layer, the second layer is the criterion layer, the third layer is the index layer, and the index layer is divided into the first-level index and the second-level index. Based on the in-depth analysis of the characteristics of entrepreneurs, this paper concludes that the evaluation points of entrepreneurs are entrepreneur's ability level, entrepreneur's credit quality and entrepreneur's property capital. The evaluation index system of entrepreneurs is shown in Table 1.
Table 1: Entrepreneur evaluation index system

<table>
<thead>
<tr>
<th>Target layer A</th>
<th>Criterion layer $B_i$</th>
<th>First-level index layer $C_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Entrepreneur character $B_1$</td>
<td>Honesty level $C_1$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Credit record $C_2$</td>
</tr>
<tr>
<td></td>
<td>Entrepreneur ability $B_2$</td>
<td>Debt paying ability $C_3$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Performance ability $C_4$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Profitability $C_5$</td>
</tr>
<tr>
<td></td>
<td>Entrepreneur property $B_3$</td>
<td>Movable property $C_6$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Real estate $C_7$</td>
</tr>
</tbody>
</table>

After building a good evaluation system, it is necessary to build a judgment matrix. The factors at the same level are compared with the criteria on the upper level, and the factors in the series are compared at the same level. For example, in this paper, the target layer A is compared with the criterion layer B, and the criterion layer B is compared with the first-level index layer C, etc. According to the result of the matrix, judge the relative importance of the element to the element at the next higher level. According to the matrix results, the relative weights under different criteria can be calculated. For example, the formula for the geometric average of each row in the matrix is shown in Formula (1).

$$\bar{w}_i = n \prod_{j=1}^{n} w_{ij}, i = 1,2,3...n$$  \hspace{1cm} (1)

We can get $\bar{W}_i = (\bar{w}_1, \bar{w}_2, \bar{w}_3...\bar{w}_n)^T$ from the above formula, and then the relative weight of each index element can be obtained by normalizing the calculation of $\bar{w}_i$, which is also the approximate value of the required feature vector. In this process, the maximum eigenvalue of the established index matrix can also be calculated. The calculation formula is shown in Formula (2).

$$\lambda_{max} = \sum_{i=1}^{n} (\bar{AW})_{ii}$$  \hspace{1cm} (2)

After calculating the maximum eigenvalues and eigenvectors of the judgment matrix, it is necessary to check the consistency of weights and combinations, and sort the test results. Only after passing the test can normalization be carried out; otherwise, the matrix needs to be rebuilt for calculation. In this process, in addition to the eigenvector and sum value corresponding to $\lambda_{max}$, it is also necessary to calculate the consistent ratio index. When the index value $C.R < 0.1$, it is proved that the system can meet the requirements. The key parameter of consistency test is C.I. The formula of C.I is $C.I = (\lambda_{max}-n)/(n-1)$, and the ratio is $C.R = C.I / (R.I.C.I)$.

The smaller the C.R, the better the consistency of the system calculation. According to the calculation, the weights of the first-level indexes are as shown in Figure 4, $B_1=0.454$, $B_2=0.272$ and $B_3=0.274$, while the comprehensive weights of the second-level indexes are $C_1=0.418$, $C_2=0.263$, $C_3=0.301$, $C_4=0.296$, $C_5=0.315$ and $C_6 = 0$. According to the
corresponding comprehensive weight, the consistency result is 
0.454×0.0723+0.274×0.92+...+0.274×0.52 = 0.01638 < 0.1, which can be inferred that the consistency of system calculation is up to standard.

Figure 4: Weight proportion chart of first-level indicators

5 CONCLUSIONS

In this paper, the author uses AHP analytic hierarchy process to construct the evaluation and analysis model of entrepreneurs, which divides entrepreneurs into three evaluation systems: entrepreneur ability level, entrepreneur credit quality and entrepreneur property capital. By comparing and establishing judgment matrix, the weight is calculated, and the consistency test proves that the model is established. The establishment and application of this model not only provide investors with the scores of entrepreneurs, help investors to choose entrepreneurs scientifically, but also provide consumers with reference guarantee for consumption.

REFERENCES

Constructive Ideas of Comprehensive Transportation Dispatch and Emergency Command System in the New Era

Bin Wei¹, Jiandong Cao²*, Weirong Luo³

*caojiandong@catsic.com

¹Sangfor Technologies Inc. Shenzhen 518000, China
²China Academy of Transportation Science Beijing 100010, China
³Sangfor Technologies Inc. Shenzhen 518000, China

Abstract: In the new Era, comprehensive transportation dispatch and emergency command system has become a concerned subject. By sorting out business characteristics, overall architecture and construction, the article summarizes development status of comprehensive transportation dispatch and emergency command system. Proceed to the next step, the article thoroughly analyzes two major changes of comprehensive transportation dispatch and emergency command system in the new Era: management transforms from data management to data asset management, and application transforms from data statistics to data intelligence. In view of the above, the article proposes overall architecture, horizontal and vertical logical architecture for comprehensive transportation dispatch and emergency command system.

Keywords: Comprehensive Transportation, Data, Dispatch, Emergency Command.

1 INTRODUCTION

In the new era, empowering transportation development with data resources and building a comprehensive transportation data center system has become an important task in the construction of a strong transportation country. Therefore, it is necessary to conduct in-depth research and discussion on the construction ideas of comprehensive transportation data brain (also known as comprehensive transportation dispatch and emergency command system).

2 CURRENT SATUS OF DISPATCH AND EMERGENCY COMMAND SYSTEM

The construction of comprehensive transportation dispatch and emergency command system started from "Highway and Waterway Traffic Safety and Smooth and Emergency Disposal System Project" in the 12th Five-Year Plan period, and developed to "Comprehensive Transportation Dispatch and Emergency Command System, Transportation Information Resources Exchange and Sharing and Open Application Platform" in the 13th Five-Year Plan period, covering transportation modes from the initial highway and waterway to various
transportation modes such as highway, railroad, waterway, civil aviation, postal service and public transportation in key cities, and covering business departments from the initial transportation industry internal coordination to cross-industry multi-departmental coordination such as public security, emergency response, health committee, land and weather \(^{(1)}\).

### 2.1 Dispatch and Emergency Command System 1.0

In the first year of the 12th Five-Year Plan, the Ministry of Transport officially issued the "Highway and Waterway Transportation Information Technology '12th Five-Year' Plan for Development", which pointed out that the Ministry of Transport organizes the construction of four major projects, including highway and waterway traffic safety and smooth and emergency disposal system, transportation economic operation monitoring and early warning and decision analysis system, highway and waterway traffic travel information service system, highway and waterway construction and transportation market credit information service system. Among them, the highway and waterway safety and emergency disposal system can be regarded as "dispatch and emergency command system 1.0".

![Construction of dispatch and emergency command system 1.0](image)

**Figure 1** Construction of dispatch and emergency command system 1.0

#### 2.1.1 Business Characteristics

Highway and waterway traffic safety and smooth and emergency disposal system is positioned to solve three major business problems: the first is to realize the visible, measurable and controllable properties of highway and waterway infrastructure and carriers; the second is to achieve the cross-regional and cross-level information sharing and coordination and command of highway and waterway; the third is to address the emergency communication guarantee and emergency disposal coordination of highway and waterway emergencies.

#### 2.1.2 Overall Architecture Model

In terms of construction content, highway and waterway traffic safety and smooth and emergency disposal system mainly constructs national and provincial road network management system, highway network management system, road transport operation management system, waterway operation management system, waterway safety operation management system and other professional systems and transportation emergency disposal comprehensive system, supplements and improves the basic database of highway and waterway, safety and emergency database, constructs the dynamic monitoring terminal covering important highway sections, bridges, emergency rescue vehicles, emergency
transport vehicles, improves wired and wireless communication system and comprehensive communication and dispatching system, etc.

2.2 Dispatch and Emergency Command System 2.0

In the first year of the 13th Five-Year Plan, the Ministry of Transport officially issued the "13th Five-Year Plan" for the Development of Transport Informatization, which proposes to focus on "three promotion, five enhancement, two guarantee" industry informatization project. Among the "five enhancement", the Comprehensive Transportation Dispatch and Emergency Command System, Transportation Information Resources Exchange and Sharing and Open Application Platform can be regarded as a common "dispatch and emergency command system 2.0".

Figure 2 Construction of comprehensive transportation dispatch and emergency command system

Figure 3 Construction of transportation information resources exchange and sharing and open application platform

2.2.1 Business Characteristics

The Comprehensive Transportation Dispatch and Emergency Command System is positioned to improve three major capabilities: the first is to improve the macro operation control capability, and realize the tracking and monitoring of the operation status of highways, railroads, waterways, civil aviation, postal services and public transportation in key cities; the second is to improve the efficiency of emergency command and dispatch, and realize the
coordination and linkage of dispatch within the transportation system, between the Ministry of Transportation and relevant ministries, and between the Ministry of Transportation and the national bureaus under the control of the Ministry; the third is to improve the macro decision support capability, and realize the prediction and warning of traffic operation in natural disasters, the prediction and warning of traffic operation in emergencies, and the prediction and warning of traffic operation in important activities, etc. [2,3]

The Transportation Information Resources Exchange and Sharing and Open Application Platform is positioned as the main hub of industry information resource exchange and sharing, focusing on three major roles: the first is a resource platform to realize the convergence and integration of industry shared resources, unified management of shared resources, and the formation of a unified transportation shared information resource pool; the second is a transmission channel to realize the collaborative scheduling of industry information resources and provide a safe, efficient, reliable and convenient data transmission channel for cross-regional and cross-system data transmission; the third is a service window to realize the departmental docking, industry sharing and opening of transportation information resources to the outside world. In particular, the transportation information resource exchange and sharing and open application platform provides data support for the nationwide and cross-regional business collaborative operation of the comprehensive transportation dispatch and emergency command system [4-7]

2.2.2 Overall Architecture Model

In terms of construction content, the Comprehensive Transportation Dispatch and Emergency Command System mainly constructs comprehensive transportation operation dynamic management system, comprehensive transportation operation prediction and early warning system, comprehensive transportation emergency command system and other comprehensive systems, and completes emergency plans, emergency cases, emergency resources, emergency watch, prediction and early warning databases; while the Transportation Information Resources Exchange and Sharing and Open Application Platform mainly builds comprehensive systems such as information resource directory service system, data exchange and sharing and open management system, data exchange and sharing and open portal system, data quality audit system, comprehensive query system, etc., and completes the exchange and sharing library, open service library and other databases.

3 NEW REQUIREMENTS FOR DISPATCH AND EMERGENCY COMMAND SYSTEM

In the new era, building a comprehensive transportation data center system with data as the key element and core driver, and constructing a comprehensive transportation dispatch and emergency command system has become an important task for the construction of a strong transportation country.

Firstly, the data management of comprehensive traffic dispatch and emergency command system changes to asset-based. In the period of Dispatch and emergency command system 1.0 and 2.0, the industry has not yet formed an objective and scientific understanding of big data and data resources value utilization, and there exist phenomena such as blindly chasing
investment in hardware facilities and belittling data resources accumulation and value mining utilization. In the new period, data has become a factor of production alongside land, labor, capital and technology, and the country treats data as an important asset for circulation, management and operation. By focusing on the whole life cycle of data, the comprehensive transportation dispatch and emergency command system collects and stores the full amount of data, deeply promotes data sharing and convergence, encourages the opening of public information resources, and promotes the intensive integration, collaborative development, efficient utilization and network sharing of production factors with data flow to form a new resource allocation model.

Secondly, the data application of comprehensive traffic dispatch and emergency command system changes to intelligence. In the period of dispatch and emergency command system 1.0 and 2.0, data application was more concerned with what was happening at present and describing the development trend. In the new period, the data application of comprehensive traffic dispatch and emergency command system has entered the intelligent stage characterized by deep excavation and fusion application. The previous "people looking for data" is gradually changing to "data looking for people", and data applications can not only predict what may happen in the future and present the development trend of things, but also guide the industry to make correct decisions based on the current development trend of things.

4 NEW IDEAS FOR THE DISPATCH AND EMERGENCY COMMAND SYSTEM

In the new era, comprehensive transportation dispatch and emergency command system 3.0 returns to the essence of "data" and is a new infrastructure that takes data as the key element and core driver to promote the continuous integration and interaction of transportation activities in physical and virtual space [8-10]

Figure 4 Overall architecture model of dispatch and emergency command system 3.0
In terms of comprehensive applications, the focus is on cross-industry, cross-sector, cross-department and cross-level applications such as decision support and evaluation, scheduling and emergency command, government office management and services, information resource sharing and opening, network security and operation and maintenance protection, etc., to realize the deep integration of new generation information technology and transportation management and decision-making. For example, the comprehensive application of decision support and evaluation can be considered to achieve digital, graphical and panoramic control of infrastructure planning, planning, construction progress and overall operation of various modes of transportation, and to achieve scientific assessment of the development level, development speed, development quality, development direction, interaction between the industry and national economy and contribution of various modes of transportation.

In terms of data management, it will accelerate the improvement of the industry data resource chain from diversified collection, integration and sharing, open circulation to social application, establish a comprehensive transportation information resource directory covering all elements and multiple levels, build a complete and standardized comprehensive transportation industry information resource "general ledger", promote cross-regional, cross-industry, cross-sector and cross-level comprehensive transportation data resources to the dispatch and emergency command system, assess the data security risks that the comprehensive transportation dispatch and emergency command system may face, and do a good job of risk control such as sensitive data identification, sensitive data access, data flow risk monitoring and data leakage traceability.

At the same time, the portal portal, map services, information resources, basic conditions, security prevention and control, standards and specifications and other basic common architecture for unified design, unified construction, so as to form an integrated comprehensive traffic dispatch and emergency command system support and security system. For example, the infrastructure of basic conditions can be considered to realize the intensive construction and integrated management of various hardware and software support resources such as operating system and middleware, database management system, computing resources, storage resources, network resources, security resources, etc., so as to provide a safe, reliable, independent and controllable basic environment support for comprehensive application and data management.

5 CONCLUSION

By proposing overall architecture, horizontal and vertical logical architecture, the thesis provides new ideas for comprehensive transportation dispatch and emergency command system with data as the core. In view of the above, this study fills the gap in the research of construction ideas in this field.

REFERENCES


Design and Implementation of Lightweight Web Asset Identification System

Zhengde Li
lizd@qfnu.edu.cn

Qufu Normal University School of Computer Science, Qufu Normal University, China Rizhao, Shandong, China

Abstract: This paper mainly describes the purpose of Web asset identification system research and development, design ideas, implementation process and so on. Web asset identification system is very popular in the attack and defense confrontation because the client does not need to rely on the environment, and can detect the target site through a variety of methods. With the Web Asset identification system, task scripts can be executed remotely without the need for real-time client execution. Asset detection can be done online. The Web asset identification system supports active and passive scanning. Web asset identification system is mainly divided into information collection module, target site module, file management module, virtual terminal module, user management module, export test module. Each module of the Web asset identification system has its specific meaning and function, but there are connections and interactions between each module. Web asset recognition system is different from the current mainstream asset recognition system, especially the passive scanning module of the Web asset recognition system adopts the form of crawler to crawl the website, which is not available in other Web asset recognition systems.

Keywords: Penetration Test, Web Asset identification System, B/S; Node. Js, Vue.

1 INTRODUCTION

On December 27, 2021, the net letter office issued the "difference" national informatization planning, "planning" digital as the core, puts forward seven development goals, deployed 10 major task, which, in terms of safety, stressing the need to pay equal attention to security and development, in order to realize the network space will markedly enhance its capability of governance and security as the goal, We will deepen the security concept of moving forward and taking precautions in advance, strengthen the mechanism for coordinating cyber security information, develop cyber security technologies and related products, and enhance the capability of independent cyber security defense. As China's network security industry promotion policies continue to increase, the product system is gradually improved, and ecological construction continues to advance, the network security industry continues to grow.
The scale of China's cybersecurity industry reached 156.359 billion yuan in 2019, up 17.1% from 2018, according to the White Paper on China's Cybersecurity Industry (2020) released by the China Academy of Information and Communications Technology. In essence, cyberspace is an environment of interaction and collaboration between people. Network technology is only a tool to connect people together. The core of cyber security is still the issue of people. Under this cognition, it can be regarded as a way to solve new problems in the new era to look back on the wisdom of traditional Chinese culture and use the essence of traditional Chinese culture to grasp the essence of human beings from the perspective of human nature and understand the method of establishing interactive order between people. Cyber security is about you and me. One end of it is connected to Cathay And the other to Min 'an. The Internet has increasingly become a new space for people to study, work and live, a new platform for people to access public services, and an important driving force.

1) For national development. For example, nowadays, big data, artificial intelligence and other technologies are widely used, and intelligent life has penetrated deeply into people's lives. Most government affairs have been transferred to the Internet, realizing "one-stop operation". Network security risks not only involve countries and enterprises, but also involve everyone.

2) Research significance: With the rapid development of the Internet, Web applications are becoming more and more extensive. Shopping, banking, airline tickets, forums, tweets, etc. The popularization of Web application has brought great convenience to people's life, but Web security is easy to be ignored. The emergence of new Web technologies greatly speeds up the development of Web applications by enterprises and individuals, but at the same time, the vulnerabilities related to these technologies are also constantly involved in them. The research on Web security has reached the moment of thousands of catch a catch. Based on the above analysis, it is very necessary to study the principle of Web vulnerability and develop an automatic Web asset identification system. On the one hand, it can identify various vulnerabilities accurately, which greatly reduces the threat faced by the Web. On the other hand, it can send a large number of requests quickly and efficiently. Compared with manual, it is not only efficient, but also avoids human factors. On the premise of legality, collect asset fingerprints, analyze the risks of assets, mark the risks and levels of assets, notify asset vulnerabilities through correct channels, and solve problems before threats come.

3) Research status: Literature 1 puts forward the idea of using crawler to identify assets, but the disadvantage is that web fingerprint library needs to be manually added, which increases the labor cost. Reference 2 proposed the associative component discovery method based on dictionary and the associative component discovery method based on the feature of component source file, but the disadvantage is that the feature code is less. Therefore, Web asset identification system solves the above problems.

4) Research content: The main content of the research is how to identify assets. Traditional asset identification is to match HTTP response packets through the existing Web application fingerprint database. The author adopts the method of automatic collection to judge and enter the database, so as to quickly collect assets.
2 PREPARATION

Vue: Vue is a set of progressive frameworks for building user interfaces. Vue is designed to be applied layer by layer from the bottom up. Construct a data-driven Web framework through vue.js, which contains a variety of view components to realize data response and composition for developers to use, including a wide variety of component libraries for developers to choose from [3].

Node.js: Node.js is a JavaScript runtime environment based on Chrome V8 engine. Node.js uses an event-driven, non-blocking I/O model, which is not restricted by the client (browser), so that JS has the same operation permissions on files, network and operating system processes as the back end, and has little difference from the functions of Java, Python and other programming languages [3].

Python crawlers: The Requests library is a concise and simple third-party library for handling HTTP requests. Written in Python, the Library is based on URLlib and is Licensed under the E2 open source protocol. It is more convenient than URLlib, can save us a lot of work, fully meet the REQUIREMENTS of HTTP testing the biggest advantage is that the program writing process is closer to the normal URL access process. There are many ways to open a web page, the most common of which are GET and POST. GET is used to access the page by directly entering the URL in the address bar of the browser. GET () is a GET request corresponding to HTTP for obtaining HTML web pages, and data is obtained by making a request to the specified URL [4].

Database related technology: SQLite is a lightweight, open source, embedded relational database. It is an ACID-compliant relational database management system. It is a zero-configuration database that requires no configuration on the system. As an open source database, SQLite is widely used by major software companies, such as Firefox, iPhone, iPad and Android, etc [5].

With the rapid development and continuous update of the Internet, database has been more and more widely used. Database development up to now, has produced a variety of functions of different types, among them, the more widely used SQL Server database and MySQL database. MySQL database has the characteristics of small memory occupation, relatively low development cost, relatively fast running speed and can support a variety of computer programming languages, and its corresponding source code is free. Therefore, MySQL database is deeply trusted by the majority of small and medium-sized websites and corresponding enterprises [6].

This chapter explains the relevant technical basis involved in the paper. Only by understanding the basic knowledge of relevant technologies, can we continue to understand the design and principle of corresponding modules in the paper works. Web asset identification system adopts B/S architecture and MTV mode, so loose coupling relationship between components is maintained. M is primarily used to take care of business objects and database objects, T is responsible for how pages are presented to users, and V is responsible for business logic and calls M and T when appropriate.
3 SYSTEM ANALYSIS

System process analysis: The modules of the Web asset identification system can be divided into the following types: information collection module, target site module, file management module, virtual terminal module, user management module and export test module, as shown in Figure 1.

Figure 1 Overall system structure

Figure 1 Overall structure of the system System process analysis mainly analyzes information collection. Other module processes are short and are only analyzed in corresponding modules as shown in Figure 2.

Figure 2 Flow diagram of Web asset identification system

(2) Passive scan module analysis: Passive scan module Analysis book submodule includes three parts, namely the SRC platform, passive scan and active scan, among which passive scan includes. See Figure 3.
(3) Active scan module analysis: This sub-module includes three parts, namely, the SRC platform, passive scan and active scan, among which the active scan includes. See Figure 4

(4) Target site module analysis: This module includes four parts, which are adding sites, scanning sites, deleting sites, modifying sites. See Figure 5
(5) Document management module analysis: File management module analysis This module includes six parts, respectively upload, download, copy, compression, delete, rename. See Figure 6

(6) Analysis of virtual terminal module: Analysis of the virtual terminal module The virtual terminal module can remotely connect to the server on the Web and quickly enter the server to troubleshoot faults.

(7) Quick start module analysis: This module includes three parts, which are adding software path, generating registry download, running registry. See Figure 7

4 DESIGN AND IMPLEMENTATION OF THE SYSTEM

The system display module using UI is Vue. Vue has a rich library of components that can quickly complete the layout of the front end.

The development environment needs to meet the requirements of the Windows system, and also need to install Vue 2.6.11, Node.js v14.17.6, element-UI 2.15.6 and so on.
The modules in the menu include "Information Collection", "Target site", "File Management", "Virtual Terminal", "Quick start", "User Management", "Export Test" and "Panel Setting".

Click the information collection module, a new menu will pop up. The menu of information collection module includes "SRC platform for mass testing", "Active Detection scan", "Passive detection scan", and "Proxy Settings" and "environment variables" are set on the panel.

(2) Design and implementation of active scanning module:
First, the user enters the domain name of the target site; Then, click the Search button; Then, the scanning script requests information from the target site according to the domain name. The request method can be either get or post. After getting the returned information from the target site, it stores it. If necessary, the next step is comparative analysis. At the same time, the returned information is displayed on the page.

(3) Design and implementation of passive scanning module: Design and Implementation of passive scanning module The implementation process of passive scanning module is as follows: The user enters the advanced syntax of the corresponding search engine and clicks the search button. The front end sends the value entered by the user to the back end through get request. The back end calls the Python crawler script to obtain the information.

(4) Design and implementation of SRC platform for mass testing: Design and Implementation of SRC Platform for mass testing The implementation process of SRC platform module for mass testing is as follows:
Collect the mainstream emergency response center on the market on the Internet and store it in the database. When the user clicks the crawl button, the front end transmits the value to the back end through GET request to crawl the list of manufacturers and save it in the SQLite database. When the user clicks the query, the back end calls the database to query the corresponding manufacturer information.

(5) The design and implementation of the target site: the design and realization of the target site target site principle mainly through the interaction of other modules or manually add, the user can choose the site for vulnerability scanning, add site by entering a name and a domain name to achieve the effect of added manually, click ok after sent via a get request to the backend interface call for automation.

(6) Design and implementation of document management: Design and implementation of file management File management has upload, download, new, copy, compression, delete, rename, search functions

(7) Design and implementation of virtual terminal module: The benefits of using a virtual terminal are as follows: In the presence of a jumping machine environment, if you have any open web services jump machine itself, that could be deployed on the springboard for virtual terminal, then don't via SSH or RDP access to jump machine, open the browser can directly in the form of web SSH remote access network device, by detecting the in some network firewalls don't allow SSH, But it is useful in environments that allow HTTP and HTTPS. In addition, SSH client software such as putty and Secure CRT is not required.

(8) Design and implementation of fast start: the design and implementation of a fast start principle of environment variable is submitted by the user in the absolute path to automatically
generate the corresponding registry files in the background, when users click the registry to download again to download the registry file, double click on the run in the computer system registry, finally can be run in the quick launch module local procedures.

The following table shows that the Web asset identification system is superior to current asset identification system tools in terms of user customization.

### TABLE I. COMPARISON OF ADVANTAGES AND DISADVANTAGES OF WEB ASSET IDENTIFICATION SYSTEMS

<table>
<thead>
<tr>
<th>Assets system</th>
<th>Architecture</th>
<th>CDN identification</th>
<th>SRC vendor list can be crawled</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goby</td>
<td>C/S</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>lighthouse</td>
<td>B/S</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Web asset identification system</td>
<td>B/S</td>
<td>Y</td>
<td>Y</td>
</tr>
</tbody>
</table>

5 CONCLUSIONS

Web asset identification, Web vulnerability scanning, Web vulnerability utilization and other tools. They are all indispensable automatic and intelligent tools in red blue confrontation. Each tool has different advantages. For example, WebShell management tool is good at modifying the files of the target host and has high stability. Web asset identification system is a comprehensive tool, which can not only scan vulnerabilities but also collect assets. If WebShell management tools and EXP are integrated into the browser, it will be a powerful tool for individual combat. Web asset recognition system is the first successful project that the author summarizes and develops through practical experience. There are still many deficiencies in function points, such as WebShell management, EXP utilization, POC detection, etc. These function points are an indispensable part of red-blue confrontation. In the long run, WebShell management function can use the national secret to confuse the transmission data, so as to bypass the packet detection of security vendors, greatly improve the efficiency of security testing, and avoid the product interception of security vendors. EXP uses the collected Nday, 1day, 0day to conduct one-click GetShell, so as to quickly collect assets.

REFERENCES


Research on Dilemmas and Rescue Measures of Enterprises in the Yangtze River Delta Under the COVID-19 Pandemic

Lu Li¹, Wenyan Zhang¹,², Qing Liu³

21349015@suipe.edu.cn¹, *Corresponding author: 20044020@suipe.edu.cn², sjys_liuqing@163.com³

Shanghai University of International Business and Economics, Shanghai, China¹

Abstract: Due to the closure of the Shanghai area, enterprises in the Yangtze River Delta region were affected by the epidemic and faced problems such as supply chain disruption, soaring costs and lost orders. This research uses a combination of in-depth interviews and questionnaires to study the current situation, relief measures and government policies of some enterprises in the Yangtze River Delta region under the impact of the new crown epidemic, and on this basis, make relevant recommendations: (1) Operators develop "risk control" methods and rules to improve their ability to deal with unforeseen situations. (2) Companies are strengthening their innovation and exploring new models of development. (3) The government sector provides long-term, stable assistance for the development of all types of businesses, depending on the characteristics of the sector.

Keywords: COVID-19 Pandemic, Yangtze River Delta Enterprise, Rescue Measures.

1 INTRODUCTION

Since March 2022, the Yangtze River Delta city cluster, centered on Shanghai, has been affected by the epidemic to implement regional closure and control, and most enterprises' production and operation have been significantly affected. How to help industries and enterprises seek their development space and improve weaknesses is a key topic for the restoration of social and economic order and economic recovery[1].

In conjunction with General Secretary Xi Jinping's request for young students to "measure the land of the motherland with their feet and discover the spirit of China with their eyes", the researcher participated in the "Visit Ten Thousand Enterprises, Read China" social survey project by distributing questionnaires and conducting online interviews to 137 enterprises in the Yangtze River Delta region.

This paper selects the manufacturing industry, foreign trade industry and Internet service industry, which are affected by the epidemic, and takes several representative enterprises as the entry point for detailed analysis. Also, since micro and small enterprises are a major driver of sustainable and social development in developing countries[2], the researcher also focused on some of the smaller companies in the Internet industry.
The innovations and contributions of this paper are:

- Through the interview form of investigation, researchers can get first-hand information. The enterprise situation will be clearly judged and the industry insider will be deeply investigated.

- Focusing on the Yangtze River Delta region, analyzing the plight of enterprises and relief policies, which will serve as a reference or demonstration for the future development of the Yangtze River Delta and other regions.

- In the current post-epidemic era, it is important to accurately analyze the impact on enterprises, summarize the challenges faced by enterprises, and give corresponding suggestions to promote the steady development of enterprises under the normal management of the epidemic.

2 GENERAL DESCRIPTION OF BUSINESS DILEMMAS

The statistical analysis of the 137 questionnaires shows that the manufacturing sector accounts for the largest proportion of the enterprises studied, reaching 56%. In terms of the types of enterprises to which they belong, there are also a significant number of enterprises with foreign trade operations, accounting for about 30% (Figure 1 and Figure 2).

![Figure 1](image_url). Industry in which the company's main business is located
In Figure 3, the majority of the 137 companies surveyed were hit by the outbreak, with 58% of them being more seriously affected.

The first and second quarters were the splitting point for the seriousness of the epidemic in Shanghai, with many companies facing a drop in orders and a shortage of raw materials due to the closure of the epidemic. As can be seen from Figure 4, orders during the epidemic were basically flat or down 5% to 20% for most companies compared to the same period last year, with companies being significantly more affected in the second quarter than in the first quarter.
As can be seen in Figure 5, 74% of companies had a decrease in their annual revenue targets, which shows the serious impact of the epidemic. However, 13% of these companies still had increased revenue targets for the year. According to the survey, it is likely that the medical device industry has seen an increase in revenue due to the need for medical devices to support the epidemic.

In this paper, the automobile manufacturing industry and the automobile accessories foreign trade industry were selected as the entry points for analysis, to explore the impact of the epidemic on the production and operation of enterprises.

Nowadays, the Internet service industry is developing rapidly, and a large number of small and medium-sized enterprises in the industry are in the start-up period. The epidemic has undoubtedly brought enormous pressure and challenges to them. Therefore, the researchers selected some small and medium-sized enterprises as examples for the Internet service industry to conduct research and exchange to understand the industry insider.

2.1 Automobile Manufacturing Industry

After the outbreak of the epidemic in Shanghai, the country implemented social static management for the epidemic area in order to prevent further spread of the epidemic, which followed problems in the supply chain, supply side and demand side of the manufacturing industry, causing considerable impact on the manufacturing industry.

Labor, raw materials, and capital are essential components of the manufacturing production process\(^3\). However, with the control of the social side, a large number of laborers were prevented from returning to work, which led to the stagnation of enterprise production activities. At the same time, the original supply chain of enterprises has been interrupted by the epidemic. Some of the interviewed automotive manufacturers revealed that certain irreplaceable high-end raw materials (such as chips) are still dependent on imports. However, affected by the epidemic, enterprises had to choose to clear customs from Chengdu and other places, resulting in rising procurement costs and a lengthening of the raw material supply timeline.

At the same time, the mass travel is hampered, the demand for cars is shrinking, and the product inventory keeps accumulating, occupying corporate funds, plus the rigid expenses such as interest, staff salary and rent, which directly puts a considerable pressure on the enterprise.
2.2 Automobile Accessories Foreign Trade Industry

As of 2019, the data released by the General Administration of Customs shows that there are 406000 actual private export enterprises in China, which is the largest foreign trade entity. The European Union, the United States, Australia and other countries are China's major trade partners, accounting for more than 80% of China's total import and export[^4].

The researchers learned from exchanges with foreign trade enterprises of automobile accessories that during the epidemic, some countries raised the quarantine inspection standards for export containers of enterprises, which resulted in prolonged transportation time of export goods, increased transportation costs and increased business pressure of foreign trade enterprises.

Because the product branches of the industry are relatively small, the enterprise scale is generally small, the product competition is fierce, and the cooperation is very few. Therefore, in the face of the COVID-19 outbreak and other major public health emergencies, the prevention of "group heating" cannot be well done.

In addition, the Canton Fair, Import Expo and other large expo technology exchange sessions cannot be held as scheduled, customers cannot come to the factory to communicate with the company, making some of the main business to view products (such as car fragrance) for the automotive foreign trade enterprises product development has been greatly restricted, the production and operation of enterprises is deteriorating.

2.3 Internet Service Industry

With the booming economy in China, small and medium-sized enterprises are mushrooming, not only promoting our economic development and technological growth, but also increasing more jobs. In particular, the Internet service industry has become the choice of many investors due to its low start-up capital and small operating costs.

However, as some small and medium-sized enterprises are still in the initial stage, they are faced with difficulties in sales, competition, capital shortage and other problems, and the impact of the new crown epidemic is even more difficult. The following is an analysis of the current situation of some Internet service enterprises in the Yangtze River Delta by means of interviews.

2.3.1 Shuyanyu (Shanghai) Science and Technology Co., Ltd.

Shuyanyu (Shanghai) Science and Technology Co., Ltd. is mainly engaged in medical device product evaluation services. Affected by the epidemic, the company can only implement the home office. Due to the short establishment of the company, members do not cooperate enough, can not complete the output of data reports and project arrangements in a timely and efficient manner.

At the same time, it is difficult for the government to give the specific time of the estimated clearance and release of quarantine, thus causing negative impact on the contract performance of enterprises and customers, resulting in enterprises afraid to take orders, thus leading to a decline in order volume.
2.3.2 Shanghai Qipin Talent Technology Group

Shanghai Qipin Talent Technology Group is mainly engaged in human resources services and Internet data services.

However, the persistence of the epidemic has led to the communication between enterprises and customers being limited to online, which has reduced customer stickiness. At the same time, online office is completely dependent on the autonomy of employees, communication between colleagues is also significantly reduced, and there is also a decline in employee efficiency.

Some offline communication and promotion activities scheduled by enterprises are also forced to be transferred to online, which leads to the problem of applicants' lack of understanding and trust in corporate positions and companies when recruiting talents, and reduces the efficiency of talent recruitment.

2.3.3 Shanghai Yuwen Education Technology Co., Ltd

Shanghai Yuwen Education Technology Co., Ltd. is mainly engaged in sand painting art. Before the epidemic, the company not only had offline teaching courses, holiday parent-child activities, but also had project cooperation with international schools, providing offline teaching courses for schools, however, affected by the epidemic offline courses turned to online teaching, the effect of the courses became worse, thus affecting the publicity and promotion of business. At the same time, affected by the national double reduction policy, business negotiation cannot be handled, and enterprise financing is difficult.

3 RESCUE MEASURES OF ENTERPRISES

Although the epidemic continues to affect our country and the global economy, it is undeniably an opportunity to force companies to reflect on their own situation, how to nurture new opportunities in the crisis and turn around the unfavorable situation to maintain operations.

3.1 Broaden Channels and Expand Extension

"Raw materials" are the basis of productive enterprises. Under the influence of the epidemic, a considerable number of manufacturing enterprises were forced to stop production because of the limited supply of raw materials, which is equivalent to the end of manufacturing at the beginning. Enterprises broaden the supply channels of raw materials and timely adjust the original production strategy in case of the major public emergencies, thus reducing the losses. For example, some enterprises with intelligent equipment began to produce anti epidemic materials such as masks and protective clothing, which not only met the market demand, but also created economic value for enterprises[5].

3.2 Export to Domestic Market and Expand Domestic Market

With the rampant spread of foreign epidemics, the demand of foreign markets has declined, and with trade barriers, the profits of export goods have been maliciously suppressed. On the other hand, the epidemic situation is well controlled at home, so foreign trade enterprises can develop the domestic market. After all, the domestic population base is large and the market demand is huge. On June 22, 2020, the General Office of the State Council also issued the Implementation
Opinions on Supporting the Transfer of Export Products to Domestic Sales[6]. The policy encourages foreign trade enterprises to develop the foreign market while opening up the domestic market, which can not only help foreign trade enterprises to open up a new situation, but also enrich the supply of the domestic market. We will vigorously promote both domestic and international circulation.

3.3 Organic Combination of "Internet+" to Improve "Soft Management"

The business model of "Internet+production" has outstanding advantages in the epidemic era. In addition to the products themselves, manufacturing and foreign trade industries also involve many other functional links, such as product design, production and quality inspection.

In the post-epidemic era, however, companies are gradually discovering that simply making manufacturing smart doesn't seem to work efficiently. Therefore, developing and building a new enterprise management platform to enhance internal communication and collaboration among employees through online meetings and live streaming may be one of the important initiatives to effectively improve work efficiency in the post-epidemic era.

4 RESPONSE MEASURES OF GOVERNMENT

In the post-epidemic era, how the government can help enterprises resume work and production and effectively reduce the economic losses incurred during the epidemic closure and control has become the focus of attention of enterprises in the Yangtze River Delta region.

4.1 Reduction and Exemption of Subsidy Rigid Expenditure

In February 2020, the executive meeting of The State Council decided to exempt social insurance premiums for enterprises in stages, and enterprises suspended the payment of housing provident fund, which greatly alleviated the pressure on enterprises.

In March this year, the SASAC issued the Notice on Rent Reduction and Waiver for Small and Micro Enterprises and Individual Businesses in the Service Industry in 2022 to reduce the rent for the relevant enterprises for six months in the current year, relieving their urgent needs. The Shanghai Action Plan for Accelerating Economic Recovery and Revitalization, released by the city on May 29, also provides subsidies for companies.

4.2 Optimize Services and Promote the Resumption of Production

In order to improve the convenience of enterprises' resumption of work and production, the General Office of the State Council issued a notice requiring governments to streamline the approval and conditions for resumption of work, and set up special columns on the national government affairs platform to improve the efficiency of enterprise business approval.

In the interview with Kunshan Ansenda Microwave Technology Co., LTD., the researcher learned that the Kunshan government launched the "Kunrui" enterprise service platform, which can automatically calculate the amount of relief for enterprises, making it easier to provide relief services for enterprises. The platform has improved the efficiency of the government's services and better helped enterprises to overcome the difficulties of the epidemic.
5 CONCLUSIONS AND RECOMMENDATIONS

The epidemic sealing and control in the Shanghai region has caused various types of enterprises in the Yangtze River Delta to face different dilemmas and challenges. In general, manufacturing, foreign trade and some small and medium-sized enterprises are more affected.

The epidemic lockdown has caused supply and demand problems for manufacturing enterprises, such as increased procurement costs and stagnant production activities, which directly affected the normal production and operation of enterprises. For foreign trade industry, enterprises are limited by export control, operating costs have risen significantly. For small and medium-sized enterprises in the Internet service industry, the problems of the start-up period have been magnified under the impact of the pandemic.

The crisis has also forced companies to act quickly, flexibly adjusting their production strategies, expanding domestic markets and accelerating the digitalization of management, to unleash resilience in the midst of adversity. At the same time, the government has also introduced relevant rescue policies, such as rent reduction and exemption, subsidies and optimization of relevant handling services, to help enterprises accelerate the resumption of work and production and effectively reduce losses.

In the post-epidemic era, the researcher makes the following recommendations:

5.1 Improve the "Risk Control System"

In order to prevent major public health emergencies from hitting the pause button on production and operation, operators can develop rules and regulations to prevent and resolve major emergencies based on their own situation, such as setting up relevant "risk control" departments and strengthening daily simulation training, so as to improve the ability of enterprises to cope with emergencies.

According to the dynamic capacity theory, MSEs need to begin to ensure business continuity in unpredictable environments through building capabilities to perceive threats of opportunity, seize opportunities, and shift paradigms. This rule applies to other types of businesses as well. And the perfect risk control system is the key to shift paradigms.

5.2 Strengthen Innovation and Explore New Models of Development

In the post-epidemic era, innovation is the key to business survival. Only by accelerating the digitalization process of enterprise management, creating a new management mode, and constantly exploring and improving the business strategy of enterprises can enterprises reshape themselves, seize opportunities and continue to develop in difficulties.

For example, during the period of closure and control, enterprises can arrange online training to enhance the professional skills of employees and inject vitality into business operation; innovate management systems and try new management models to fundamentally improve the overall quality of enterprises.
5.3 Strengthen Policy Support

In the current economic environment, government departments should provide more assistance to the development of various types of enterprises, according to the characteristics of the industry. For large enterprises, the government can enhance regional talent introduction programs and improve talent recruitment subsidies and enterprise development guidance mechanisms. For MSEs, the government should provide continuous policy support, such as subsidies, tax reduction and fee reduction, to provide reliable guarantee for their survival and development. Besides, for MSEs digital transformation, the government can play a prominent role by raising digital transformation awareness, increasing labor-power competence, providing technical and financial support, and strengthening data communication infrastructure[8]. At the same time, good policies will encourage social innovation and entrepreneurship, and can further boost the country's economy.

REFERENCES

[1] Zhao Ailing, How to Offer Relief for Small and Medium-sized Enterprises Amid the COVID-19 Epidemic, China’s Foreign Trade, 2020,(03)
Discussion on the Development Strategy of Digital Marketing of Agricultural Products from the Perspective of Long Tail Theory
Take Agricultural Products with Geographical Indications as an Example

Huiru Wang¹, Xiaoyuan Yu²,*

abcwhr2016@163.com¹, *Corresponding author: xiaoyuan.yu@cgt.bitzh.edu.cn²
the faculty of education the university of Hong Kong, Kong Kong, China¹
College of Global Talents Beijing Institute of Technology Zhuhai, Zhuhai, China²

Abstract: Innovative digital marketing is an important means to improve the circulation of agricultural products. Based on the long tail theory, this paper qualitatively analyzes the development strategy of digital marketing of agricultural products in China and examines the economic value of digital marketing, taking agricultural products with geographical indications as an example. The research results show that building agricultural product IP traffic, attracting customers through all channels and experiential marketing are important links of agricultural product digital marketing. The empirical results show that digital marketing can reduce the circulation cost of agricultural products, form a price competitive advantage, and thus improve market sales. This economic value is more and more obvious in regions with higher digital marketing level and stronger policy concern.

Keywords: Long Tail Theory, Digital Marketing, Geographical Indication, Agricultural Products.

1 INTRODUCTION

In recent years, the problem of unsalable agricultural products in China is not uncommon. How to solve the problem of circulation of agricultural products in China through science and technology and innovative marketing methods has become an important topic in the current agricultural development [1]. On the whole, in the era of digital economy, China's agricultural product circulation industry is also facing the pressure of digital transformation. Due to the lack of digital marketing means, the cost of agricultural product circulation remains high, and the problem of agricultural product circulation cannot be effectively solved. Branded+digital agricultural product marketing strategy is considered to be an important means to activate the digital sales of agricultural products, and is also the main direction of modern agricultural construction in China. In the face of the increasingly diversified demand for agricultural products in the consumer market, it has become a general trend to rely on digital platforms to cultivate agricultural product brands and carry out digital marketing [2].
2 MATERIALS AND METHODS

With the gradual spread of the global epidemic and the rapid growth of online consumption, agricultural product operators and agricultural product distribution enterprises are required to take the initiative to innovate and accelerate the digital process to enhance online marketing capabilities. Huttunen (2019) [3] believed that the Internet based and networked development of agricultural products is the main direction of future agricultural development. Through digital marketing, we can use technical means to provide continuous services for the entire agricultural industry chain, which is more conducive to the cultivation and promotion of agricultural product brands.

The so-called long tail theory refers to solving the problem of high information communication costs and commodity circulation costs through Internet digital marketing in the Internet era. In the Internet era, digital marketing can accurately match more personalized market demands, which are also known as the "long tail market". Taeuschker (2019) [4] believed that the long tail market mainly refers to the digital industry market in the network era. As it can meet more and more personalized market demands, it can obtain marginal benefits, which constitutes the realistic basis for the long-term existence of digital marketing. Geissinger et al. (2020) [5] pointed out in his research that, from the perspective of the long tail theory, if digital marketing wants to succeed, it must have the characteristics of both low cost and high income, and at the same time conform to the general law of the development of the digital industry.

It is of great significance for the diversified development of China's consumer market to promote consumption growth. The so-called long tail theory refers to solving the problem of high information communication costs and commodity circulation costs through Internet digital marketing in the Internet era. In the Internet era, digital marketing can accurately match more personalized market demands, which are also known as the "long tail market". Qian et al. [6] believes that the long tail market mainly refers to the digital industry market in the network era. As it can meet more and more personalized market demands, it can obtain marginal benefits, which constitutes the realistic basis for the long-term existence of digital marketing. Lew (2008) [7] pointed out in his research that, from the perspective of the long tail theory, if digital marketing wants to succeed, it must have the characteristics of both low cost and high income, and at the same time conform to the general law of the development of the digital industry. Coelho & Mendes (2019) [8] pointed out that the popularity and development of the Internet have changed the channels for people's information sharing and data dissemination, enabling agricultural product marketing information to be spread on the Internet platform at a lower cost. It is precisely because of the convenience of this information circulation that provides a prerequisite for the development of the long tail market.

2.1 Create Agricultural Product IP Traffic

From the perspective of the long tail theory, the success of digital marketing of agricultural products should first ensure the stability of the customer source, that is, ensure the continuous attention of the target consumer groups [9]. In the era of mobile Internet, the creation of agricultural products IP pays more attention to the publicity of content. The cost of capital and time invested in the early stage is relatively low, and the threshold for publicity is not high, which creates favorable conditions for attracting agricultural products IP traffic. At this stage,
the primary reason for the poor circulation of agricultural products between urban and rural areas in China is the separation between producers and consumers.

At the same time, based on the existing agricultural products, extend the industrial chain to form new services such as farm experience and package combination. For end consumers, through digital marketing, they can effectively obtain agricultural product information and product traceability, which can improve their trust in agricultural product quality, improve their recognition of agricultural product brand value in the long-term interaction process, and finally form a sustainable consumption behavior.

2.2 Attract Customers Through All Channels

With the continuous development of the Internet, consumers have more and more access to information, which creates favorable external conditions for the establishment of digital marketing channels for agricultural products. In the perspective of the long tail theory, digital marketing of agricultural products must continuously attract users through innovative marketing channels. Based on the Internet platform, the marketing information is implanted in the PC and mobile terminals to attract potential consumers with low information dissemination costs, and the defect of insufficient online experience is supplemented through the opening of offline stores. For this reason, based on the existing mobile media means in China, digital marketing of agricultural products mainly attracts customers from the following channels: first, short video. Through short video apps such as Tiktok and Kwai, farmers can directly provide consumers with detailed information about agricultural products by video transmission, while attracting customers by personalized publicity means. The second is WeChat applet, which is used to help farmers by entrusting a third party to build a WeChat applet platform, providing product price information to directly attract end customers to buy. The third is all kinds of fresh food apps, mainly agricultural product sales platforms built by circulation enterprises such as Hema Fresh Food and Pupu Supermarket, which attract customers based on early marketing methods to achieve the purpose of digital marketing. Fourth, it is common in all kinds of micro malls, similar to Taobao online stores, but with more extensive publicity means.

3 RESULTS

3.1 Determination of Research Object

Based on the existing theoretical analysis, this paper will try to test the economic value of digital marketing of agricultural products through empirical analysis. As mentioned above, the digital marketing of agricultural products in China is at the initial stage, and the number of agricultural products involved in digital marketing is relatively small. Therefore, this paper takes geographical indication agricultural products as the research object. The main reasons for choosing agricultural products with geographical indications are as follows: on the one hand, the selection criteria for agricultural products with geographical indications are relatively strict, most of which are agricultural products with local characteristics in China, with high quality, high economic value and high recognition in the consumer market, so they can meet the consumption demand of the long tail market. On the other hand, agricultural products with geographical indications are highly concerned by local governments, easy to obtain support from social capital, and have the advantage of large-scale operation, which is also the main
direction of the future circulation of agricultural products in China. The study of agricultural products with geographical indications has a good policy reference value. Based on this, this paper will take China's geographical indication agricultural product catalog as an example to empirically test the enabling effect of digital marketing on the circulation process of geographical indication agricultural products.

3.2 Data Selection

The interpreted variable. The purpose of digital marketing of agricultural products is to accelerate the turnover of agricultural products and realize the jump from production to consumption. Therefore, this paper will reflect the results of the circulation of agricultural products from three aspects: price, sales and circulation cost. The market price is obtained from various digital platforms, and the sales volume (annual sales volume) and circulation cost (proportion of circulation cost in sales volume) are obtained through questionnaire.

Core explanatory variables. As mentioned above, digital marketing has gone through three aspects: creating IP traffic, attracting customers through all channels, and experiential marketing. Therefore, this paper uses the Likert 10 point method to comprehensively score each link of digital marketing of agricultural products with geographical indications, and summarizes the scores to measure their digital marketing level.

Control variables. In order to make the empirical regression results more accurate, this paper also controls the following variables: digital marketing cycle, which is expressed by the starting time of digital marketing. The registration time is expressed by the time of entering the agricultural product catalog of geographical indications. Location: 1 for the eastern region and 0 for the rest. Lagging items of price, sales volume and circulation cost. Frequency of interaction with end customers. Government support is expressed by the annual public promotion frequency of government departments.

This paper selects 149 agricultural products with geographical indications based on China's agricultural product catalog of geographical indications in 2021, obtains relevant data by issuing online questionnaires and expert scoring methods, and revisits some samples with obvious errors in the questionnaire data. If the data quality still does not meet expectations, delete them, and finally obtain 138 data sets related to agricultural products with geographical indications. The relevant data and descriptive statistics involved in this paper are detailed in Table 1. This paper conducts logarithmic processing on price, sales data and interaction frequency. From the perspective of market price and sales volume variables, there are obvious differences between market price and sales volume of various kinds of agricultural products with geographical indications because of the large audience differences. From the perspective of circulation cost, the largest proportion of circulation cost exceeds 50%, and the overall average value also exceeds 27%, which shows that the circulation cost of agricultural products with geographical indications in China is too high, which is not conducive to the efficient marketing of agricultural products in China. From the perspective of digital marketing level, there are also significant differences in the performance of each sample, which provides a basis for this empirical study. In addition, in the survey, this paper found that the longer the registration time of agricultural products with geographical indications, the higher the level of digital marketing. Since 2018, with the gradual rise of short video platforms, the marketing
methods and models of agricultural products with geographical indications have been greatly enriched, providing important digital media for the marketing of agricultural products in China.

4 MODEL DESIGN

In order to test the effect of digital marketing of agricultural products with geographical indications, this paper constructs the following multiple dynamic regression model, as shown in Formula (1) - (3):

\[
\text{Price}_t = \alpha + \alpha_1 \text{Digit}_t + \beta_1 \text{Price}_{t-1} + \lambda_i X_i + \varepsilon \\
\text{Sale}_t = \alpha + \alpha_2 \text{Digit}_t + \beta_2 \text{Sale}_{t-1} + \lambda_i X_i + \varepsilon \\
\text{Cost}_t = \alpha + \alpha_3 \text{Digit}_t + \beta_3 \text{Cost}_{t-1} + \lambda_i X_i + \varepsilon
\]  

(1) \hspace{1cm} (2) \hspace{1cm} (3)

Considering the wide coverage of agricultural products with geographical indications in China, and the degree of market concern for agricultural products with different geographical indications, their prices, circulation costs and sales may not be comparable. Therefore, in order to reduce the error of the regression results, this paper adds a lag term of the explained variable in equation (1) and equation (3). \(X_i\) is the set of control variables in this paper, \(\varepsilon\) is an error item.

4.1 Benchmark Regression Analysis

Table 1 shows the benchmark regression results of this paper. From the fitting results of lag variables, the market price, sales volume and circulation cost of agricultural products with geographical indications in the current period are significantly positively related to the previous period. Therefore, if this continuity feature is not considered, it may lead to errors in the model fitting results. The fitting coefficients of the core explanatory variable \(\text{Digit}\) and the price, sales and circulation cost of agricultural products are -0.072, 0.108 and -0.317, respectively, and all pass the significance test at the 1% level, which indicates that digital marketing of agricultural products is conducive to reducing their market price and circulation cost, increasing their market sales and enhancing their market competitiveness. In this paper, the author believes that digital marketing of agricultural products is conducive to reducing the circulation of agricultural products, allowing producers to face the end consumers directly. The reduction of circulation costs can bring more price competitive advantages to geographical indication agricultural products, thus promoting the growth of their market sales.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Price</th>
<th>Sale</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimation coefficient</td>
<td>T statistics</td>
<td>Estimation coefficient</td>
</tr>
<tr>
<td>Price-1</td>
<td>0.725***</td>
<td>4.821</td>
<td></td>
</tr>
<tr>
<td>Sale-1</td>
<td>0.609***</td>
<td>8.273</td>
<td></td>
</tr>
<tr>
<td>Cost-1</td>
<td></td>
<td></td>
<td>0.711***</td>
</tr>
</tbody>
</table>
Dig -0.072*** -11.829 0.108*** 6.092 -0.317*** -15.026
Period -0.023*** -13.209 0.007** 2.447 -0.104* -1.821
Time 0.217** 2.317 0.298*** 4.098 -0.215 -0.038
Location 0.108*** 4.092 0.167*** 3.821 -0.117*** -4.499
Interact 0.015*** 1.997 0.031** 2.147 0.203 0.102
Gov 0.115*** 4.821 0.043*** 3.094 0.107 0.03
R2 0.516 0.572 0.552 0.572 0.572 0.572

Note: * * * means passing the significance test at 1% level, * * means passing the significance test at 5% level, * means passing the significance test at 10% level, the same below.

4.2 Differentiation Test

Furthermore, based on the variable of digital marketing level, this paper divides the samples into low level samples of digital marketing and high level samples of digital marketing, and tests them by samples. As shown in Table 2, in the low level of digital marketing, digital marketing has been proved to only reduce circulation costs and market prices, and has no significant effect on sales. Compared with the sample with high digital marketing level, the economic value of digital marketing in the sample with low digital marketing level is limited. This paper believes that the higher the level of digital marketing, the more conducive to grasping the consumption demand of the long tail market, thus forming a profit increment and bringing more significant economic benefits. In addition, considering that many local governments have participated in the market promotion of agricultural products with geographical indications, the market reputation of agricultural products with geographical indications has been improved to a certain extent.

TABLE II. DIFFERENTIAL REGRESSION RESULTS (BY DIGITAL MARKETING LEVEL)

<table>
<thead>
<tr>
<th></th>
<th>High digital marketing level</th>
<th>Low level of digital marketing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dig</td>
<td>-0.098*** (-9.273)</td>
<td>-0.404*** (-6.093)</td>
</tr>
<tr>
<td></td>
<td>0.127*** (7.726)</td>
<td>-0.023* (-1.790)</td>
</tr>
<tr>
<td>Other variables</td>
<td>Control Control Control</td>
<td>Control Control Control</td>
</tr>
<tr>
<td>Intercept term</td>
<td>2.176*** (12.266)</td>
<td>21.061*** (14.291)</td>
</tr>
<tr>
<td></td>
<td>14.017*** (12.815)</td>
<td>14.926*** (14.016)</td>
</tr>
<tr>
<td>R2</td>
<td>0.</td>
<td>0.</td>
</tr>
</tbody>
</table>

5 CONCLUSION

Innovating digital marketing means is an important direction to promote the circulation of agricultural products. Based on the perspective of the long tail theory, taking the agricultural products with geographical indications as an example, this paper qualitatively analyzes the
development strategy of digital marketing of agricultural products in China, and empirically tests the economic value of digital marketing. The research results show that building agricultural product IP circulation, attracting customers through all channels and experiential marketing are important means of agricultural product digital marketing. The empirical results show that digital marketing can improve the market sales by reducing the circulation cost of agricultural products and forming price competitive advantages. The results of differential regression show that the higher the level of digital marketing, the stronger the marginal positive impact on the marketing of geographical indication agricultural products, and policy support plays an important role in it. Based on this, in the context of the digital economy, to innovate the digital marketing mode of agricultural products and make full use of the network platform to promote agricultural products, local governments also need to strengthen the guidance and support of digital marketing of agricultural products, so as to provide assistance for agricultural products to move towards the consumer market.

REFERENCES

Research on the Relationship between Innovation Capability and Corporate Performance: The Moderating Effect of Value Cognition Complexity and R&D Leap

Litian Chen¹,a, Xinran Chang¹,b*, Junhong Zhong¹,c, Yufeng Wang¹,d
daisy0123@126.com¹, *2028013206@qq.comb, 1579747581@qq.comc, 493125335@qq.comd
School of Business Administration, Zhejiang Gongshang University, Hangzhou, Zhejiang, China¹

Abstract: Based on the theories of value cognition and punctuated equilibrium, this paper takes the panel data of 73 listed companies of AI concept stocks in China from 2011 to 2021 as samples. It analyses and tests the impact of innovation capability on corporate performance, the moderating effect of value cognition complexity during this process, and the activating effect of R&D leap on such effect. The research shows that corporate innovation capability makes a positive impact on corporate performance; value cognition complexity negatively regulates the positive relationship between innovation capability and corporate performance and the relationship between innovation capability and corporate performance is jointly regulated by the R&D leap and value cognition complexity. This research is of important theoretical and practical significance for exploring the improvement in corporate innovation capability and the remodelling of the cognitive framework.

Keywords: Innovation Capability, Value Cognition, R&D Leap.

1 INTRODUCTION

With the rapid development and application of emerging technologies, such as artificial intelligence (AI) and big data, these technologies not only inject new momentum into the global economy but also bring about a disruptive revolution to the development of enterprises [13, 16]. Its strong autonomy and deeper learning and decision-making capabilities help enterprises to design intelligent products, novel services and invent new business models and organizational models, thus creating value. However, due to the rapid updating of emerging technologies, in the face of the transformation of emerging technologies involving AI, enterprises need to have a strong innovation capability and innovation output to survive and develop amid fierce competition. So, how the enterprises in the wave of AI technology choose an appropriate way for improving their innovation capability based on the value cognition structure becomes the key to gaining a competitive edge. In fact, there remains a research gap in the existing literature.

In recent years, scholars focusing on strategic management have started to emphasize the role of managers’ cognitive factors in the establishment of dynamic capabilities to explain the path creation or punctuation in the dynamic development of organizational capabilities [29]. However,
capability research and management cognition research in the strategic management field has been developing on two parallel tracks [6]. Essentially, the improvement in innovation capability is an organizational learning process that contains knowledge search and selection, which is affected by the cognitive framework of enterprises [20]. On the one hand, the information processing theory holds that an enterprise with a complex, decentralized cognitive framework has more complete sources of knowledge so that it can more accurately identify and select opportunities in the environment and improve the quality of risk decision-making [2, 5, 20]. On the other hand, the social classification theory believes that an enterprise with a complex, decentralized cognitive framework will see a longer psychological distance in the decision-making process, increasing the cost of making risky decisions [10]. Thus, there is a controversy on the role of value cognition in the improvement of innovation capability and the growth of corporate performance.

Also, the R&D mode is important for the continuous improvement of enterprises’ innovation output and competitiveness. Some studies believe that it is not necessary for enterprises to match the level of high value cognition in the case of a small R&D leap due to the small risk of innovation [24]. But, some studies argue that enterprises need to match higher value cognition in the case of a small R&D leap to achieve the potential capabilities necessary for improving their innovation capability [1]. Thus, the impact of R&D leap and value cognition on corporate innovation under different matching scenarios is still controversial.

Overall, Based on the theories of value cognition and punctuated equilibrium, this paper takes the panel data of 73 listed companies of AI concept stocks in China from 2011 to 2021 as samples. It analyses and tests the impact of innovation capability on corporate performance, the moderating effect of value cognition complexity during this process, and the activating effect of R&D leap on such effect.

2 RESEARCH HYPOTHESES

2.1 Innovation Capability and Corporate Performance

According to the resource-based theory, the outstanding innovation capability of an enterprise, as the main driving force for its advantage in market competition, will also directly act on the enterprise’s innovation performance [7]. First, a strong innovation capability of an enterprise can help it effectively digest and absorb knowledge related to scientific and technological production and convert the scientific value and technological value into economic value, thus improving its innovation performance [31]. Second, the innovation capability of an enterprise can make it favored by external investors as these investors believe that innovative products can bring more benefits; and investors’ confidence in the enterprise’s innovation capability has contributed to the increase in their capital investment in the enterprise and provided some financial guarantee for the improvement in the enterprise’s innovation performance [17]. Based on the above, this paper holds that there exists a positive correlation between the innovation capability of an enterprise and corporate performance and puts forward Hypothesis 1 as follows:

H1: Corporate innovation capability makes a positive impact on corporate performance.
2.2 The Impact of Value Cognition Complexity on the Relationship between Innovation Capability and Corporate Performance

Complexity refers to the breadth of knowledge contained in the knowledge structure of managers. It combines the links of value chain creation, i.e. the number of value creation links that managers pay attention to [3]. An enterprise with a complex, de-centralized cognitive framework will have a longer psychological distance in the decision-making process, thus increasing the cost of risk decision-making [10]. Moreover, emerging technologies are complex and highly uncertain, which increases the difficulty of enterprises’ technology R&D [8]. As a result, enterprises are forced to pay attention to more external stimuli and react to them. At this point, a complex, decentralized cognitive framework will make it more difficult to achieve recognition among decision-making sub-groups [10, 18]. The lack of identity will result in prejudice and the stereotype effect among different decision-making sub-groups, and increase task and relationship conflicts among teams [12]. Based on the above, Hypothesis 2 is proposed as follows:

H2: Value cognition complexity negatively regulates the positive relationship between innovation capability and corporate performance. This means the positive relationship between corporate innovation capability and corporate performance will be weakened in the case of higher value cognition complexity.

2.3 Joint Interaction of R&D Leap and Value Cognition Complexity on the Relationship between Innovation Capability and Corporate Performance

Scholars Mudambi et al. (2011) used for the first time in 2014 the concept of “R&D leap” to describe the dynamic alternating process for exploratory innovation and utilization innovation. The transformation of innovation is a dramatic change that requires all elements in the enterprise to change their practices [19]. A higher degree of complexity forces more enterprises to pay attention to and respond to external stimuli [11, 26]. If too much attention is paid to the value link, high sunk costs will be incurred in the reallocation of resources [3], thus increasing the risk and costs of technological innovation. Conversely, in the case of low complexity, managers pay less attention to the value link. During the transformation, enterprises need to change less inertia such as systems, so they are highly flexible and spend less sunk costs [3], which is conducive to their technological innovation. Based on the above, Hypothesis 3 is proposed as follows:

H3: The relationship between innovation capability and corporate performance is jointly regulated by R&D leap and value cognition complexity. In the case of small R&D leap and low value cognition complexity, this relationship is positive.

Figure 1: Conceptual framework
3 METHODS AND MATERIALS

3.1 Samples and Data Sources

This paper studies the listed companies of Flush AI concept stocks from 2011 to 2021 as a case. The reason for such a choice is as follows: (1) Highly reliable data. (2) Highly matched research problems. AI is the frontier field of emerging technologies. This field features intensive R&D activities of AI enterprises and a strong R&D foundation of enterprise teams. Meanwhile, for more accurate measurement, we performed an initial screening of data by taking the following measures:

First, excluding companies with poor performance and with the marks of ST, ST* and PT, etc. Second excluding companies with missing or incomplete data; Meanwhile, considering the lagging effect of innovation capability on corporate performance, this paper takes the three-year average of corporate performance. Finally, a total of 73 enterprise samples were obtained. The annual reports and social responsibility reports were downloaded from CNINFO to manually encode data and obtain value cognition data. Patent data, R&D data, the proportion of independent directors, shareholding percentage of institutional investors, the proportion of state-owned shares, dual roles of the CEO, enterprise age, enterprise size, and other indicator data of sample enterprises were downloaded from the CSMAR Database. Also, the above data were supplemented and verified in combination with the annual reports of sample enterprises.

3.2 Variable Measurement

3.2.1 Dependent Variable: Corporate Performance

Based on Mudambi et al. (2014) Tobin’s Q value is selected as the proxy variable of corporate performance in this paper. That the value is larger than 1 indicates the enterprise can sell at a price higher than its asset cost or the enterprise is creating value.

3.2.2 Independent Variable: Innovation Capability (INN)

Compared with other data, corporate patent data is not easy to be manipulated and represents a good way to measure innovation capability \(^{[29]}\). So, the number of patents of listed companies is selected in this paper as the proxy variable of innovation capability. With reference to Hottenrott et al. (2012), the number of patents is measured by the number of enterprises’ applications for patents of invention. Also, the data are processed by referring to the calculation method of Chen et al. (2020):

\[
\text{Number of corporate patents} = 1 + \text{the natural logarithm of the total number of patent applications for the enterprise in the current year}
\]

3.2.3 Moderator Variable: Value Cognition Complexity (Nc)

By reference to the coding research design by Nadkarni, et al. (2007), and Wu (2011), the content analysis method is used in this research to describe the complexity of value cognition in the R&D process of Chinese enterprises. The steps are shown as follows:
Step 1: Statement recognition stage. The annual reports and social responsibility reports of enterprises are studied. Then, according to the coded vocabulary summarized by Wu (2011) in his research, identify and record the statements that enterprises take into account when they plan innovation strategies, as shown in each annual report.

Step 2: From the selected statements, identify the number of links in the value creation chain (such as R&D, production, market, manpower, and operation) that the enterprises consider and count it as the value of value cognition complexity. The greater the complexity value, the more value creation links are considered in the process of corporate decision-making and cognition.

The specific vocabulary is shown in Table 1

3.2.4 Moderator Variable: R&D Leap (RDL)

This paper refers to Mudambi et al. (2014) and Swift (2016), particularly their methods of measurement of R&D leap. The specific steps are shown as follows:

Step 1: Calculate the autoregressive model residual $u_{itn}$ of the ith enterprise in year t and then proceed to the modeling of the next step;

Step 2: Calculate the GARCH model residual of the ith enterprise in year t, which measures the degree to which the enterprise’s R&D expenditure in the year deviates from the predicted value that shows its historical trend;

Step 3: Then, calculate the studentized residual $e_{itn}(stud)$ for the R&D expenditure GARCH model of the ith enterprise in year t for subsequent comparative research. The specific calculating formula is shown as follows:

\[
(1) \quad e_{itn}(stud) = \frac{e_{itn}}{s_{itn} \sqrt{1 - h_{int}}} , \quad \text{where} \quad s_{itn} \text{ is the standard deviation of } e_{itn} \text{ and } h_{int} \text{ is the impact of } u_{itn} \text{ of the ith enterprise in year t on the entire estimation.}
\]

Step 4: Compare the absolute value of studentized residuals of each enterprise during the ten years from 2011 to 2020 and find the maximum value $e_t(max)$ during the observation, which is the R&D leap of the ith enterprise.

3.2.5 Control VARIABLES

In this paper, control variables based on research on innovation capability and corporate performance are selected. First, the strategic deployment of an enterprise will affect its innovation decisions. Thus, the proportion of independent directors, the shareholding percentage of institutional investors, the proportion of state-owned shares, and the dual roles of the CEO are adopted as control variables in this paper. Second, due to the impact of corporate innovation output by corporate resources [28], enterprise age, enterprise scale, and asset-liability ratio are also controlled in this paper. Finally, given the impact of the difference in the industry and the year, the innovation output on the enterprise (Chen, et al., 2018; Chen, et al., 2021), the industry and the year are also controlled here.
4 DATA RESULTS

4.1 Descriptive Statistics

In this paper, stata17.0 is used to make descriptive statistics on the data to obtain the mean and variance, and Pearson correlation coefficient is used to describe and test the correlation between variable data. The relevant results are shown in Table 2. Furthermore, all the variance expansion factors are tested. It is found that all VIF values are smaller than 3 and the mean value is 1.37, indicating that the model in this paper is free of the serious multicollinearity problem.

4.2 Hypothesis Test

In order to avoid the problems of heteroscedasticity, sequence correlation, and cross-section correlation, the Driscoll-Kraay (D-K for short) standard error is used for estimation with the aid of stata16.0. Meanwhile, the original hypotheses are rejected by the Hausman test results, so the fixed effect model is adopted. Additionally, the interaction term variable in the model is centralized to avoid multicollinearity.

(1) Testing the effect of innovation capability on corporate performance

This paper conducts a multiple regression analysis on the relationship between core variables, with the results shown in Table 3. According to Model 2, the regression coefficient of innovation capability is 0.056 (p<0.05), indicating a significant positive correlation between innovation capability and corporate performance. Thus, Hypothesis 1 is verified.

<table>
<thead>
<tr>
<th>Table 1. Coding vocabulary of value creation link</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Value creation links</strong></td>
</tr>
<tr>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td><strong>Main links</strong></td>
</tr>
<tr>
<td>R&amp;D</td>
</tr>
<tr>
<td>Design</td>
</tr>
<tr>
<td>Production</td>
</tr>
<tr>
<td>Marketing</td>
</tr>
<tr>
<td>Supply</td>
</tr>
<tr>
<td>Service</td>
</tr>
<tr>
<td><strong>Support links</strong></td>
</tr>
<tr>
<td>Information management</td>
</tr>
<tr>
<td>Relationship management</td>
</tr>
<tr>
<td>Operation management</td>
</tr>
<tr>
<td>Manpower</td>
</tr>
</tbody>
</table>
Table 2. Descriptive statistics of variables and Pearson correlation analysis

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>Perf orm</th>
<th>INN</th>
<th>Nc</th>
<th>RDL</th>
<th>Age</th>
<th>Size</th>
<th>debt</th>
<th>Inst</th>
<th>Indep</th>
<th>Stat</th>
<th>Dual</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perf orm</td>
<td>3.27</td>
<td>1.88</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INN</td>
<td>2.87</td>
<td>1.89</td>
<td>0.13</td>
<td>7***</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nc</td>
<td>4.73</td>
<td>1.85</td>
<td>0.12</td>
<td>0.23</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RDL</td>
<td>2.42</td>
<td>4.30</td>
<td>0.09</td>
<td>0.16</td>
<td>0.14</td>
<td>3***</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>2.83</td>
<td>0.30</td>
<td>0.16</td>
<td>0.20</td>
<td>0.00</td>
<td>0.16</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size</td>
<td>22.1</td>
<td>1.24</td>
<td>0.32</td>
<td>0.64</td>
<td>0.28</td>
<td>0.25</td>
<td>0.32</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>debt</td>
<td>0.33</td>
<td>0.17</td>
<td>0.33</td>
<td>0.30</td>
<td>0.29</td>
<td>0.34</td>
<td>0.19</td>
<td>0.56</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inst</td>
<td>2.86</td>
<td>6.72</td>
<td>0.04</td>
<td>0.07</td>
<td>0.05</td>
<td>0.13</td>
<td>0.28</td>
<td>0.14</td>
<td>0.11</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indep</td>
<td>0.37</td>
<td>0.05</td>
<td>0.09</td>
<td>0.10</td>
<td>0.01</td>
<td>0.19</td>
<td>0.05</td>
<td>0.09</td>
<td>0.05</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stat</td>
<td>0.01</td>
<td>0.08</td>
<td>0.00</td>
<td>0.07</td>
<td>0.04</td>
<td>0.21</td>
<td>0.02</td>
<td>0.11</td>
<td>0.03</td>
<td>0.09</td>
<td>0.14</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Dual</td>
<td>0.33</td>
<td>0.47</td>
<td>0.06</td>
<td>0.19</td>
<td>0.03</td>
<td>0.19</td>
<td>0.10</td>
<td>0.24</td>
<td>0.19</td>
<td>0.05</td>
<td>0.14</td>
<td>0.12</td>
<td>1</td>
</tr>
</tbody>
</table>

(2) Testing the moderating effect of value cognition complexity

To further test the moderating effect of value cognition complexity, this moderating variable and the corresponding product term are included in this model. According to Model 3, the regression coefficient between innovation capability and value cognition complexity is -0.013 (p<0.1), indicating that value cognition complexity negatively regulates the relationship between innovation capability and corporate performance. Thus, Hypothesis 2 is verified.

(3) Testing the moderating effect of value cognition complexity and R&D leap on innovation capability and corporate performance

To further test the regulatory effect of the dual interaction between value cognition complexity and R&D leap, the quadratic product term and cubic product term of innovation capability, value cognition complexity, and R&D leap are included in the model. According to Model (4), the regression coefficient of the cubic product term (INN-Nc-RDL) is 0.004 (P<0.001), indicating
that the interaction term between value cognition complexity and R&D leap positively affects the relationship between innovation capability and corporate performance. Thus, Hypothesis 3 is verified.

Table 3: Stratified regression analysis results

<table>
<thead>
<tr>
<th></th>
<th>Model (1)</th>
<th>Model (2)</th>
<th>Model (3)</th>
<th>Model (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-0.250</td>
<td>-0.275</td>
<td>-0.290</td>
<td>-0.254</td>
</tr>
<tr>
<td></td>
<td>(-1.31)</td>
<td>(-1.43)</td>
<td>(-1.53)</td>
<td>(-1.35)</td>
</tr>
<tr>
<td>Size</td>
<td>-0.286***</td>
<td>-0.350***</td>
<td>-0.360***</td>
<td>-0.343***</td>
</tr>
<tr>
<td></td>
<td>(-5.46)</td>
<td>(-6.18)</td>
<td>(-6.36)</td>
<td>(-6.08)</td>
</tr>
<tr>
<td>Debt</td>
<td>0.217</td>
<td>0.145</td>
<td>0.131</td>
<td>0.103</td>
</tr>
<tr>
<td></td>
<td>(0.90)</td>
<td>(0.60)</td>
<td>(0.53)</td>
<td>(0.42)</td>
</tr>
<tr>
<td>Inst</td>
<td>0.014**</td>
<td>0.014**</td>
<td>0.015**</td>
<td>0.014**</td>
</tr>
<tr>
<td></td>
<td>(2.00)</td>
<td>(1.96)</td>
<td>(2.11)</td>
<td>(2.02)</td>
</tr>
<tr>
<td>Indep</td>
<td>0.153</td>
<td>0.309</td>
<td>0.349</td>
<td>0.395</td>
</tr>
<tr>
<td></td>
<td>(0.25)</td>
<td>(0.50)</td>
<td>(0.57)</td>
<td>(0.65)</td>
</tr>
<tr>
<td>Stat</td>
<td>-0.062</td>
<td>-0.185</td>
<td>-0.255</td>
<td>-0.348</td>
</tr>
<tr>
<td></td>
<td>(-0.15)</td>
<td>(-0.46)</td>
<td>(-0.58)</td>
<td>(-0.83)</td>
</tr>
<tr>
<td>Dual</td>
<td>-0.126*</td>
<td>-0.101</td>
<td>-0.101</td>
<td>-0.100</td>
</tr>
<tr>
<td></td>
<td>(-1.77)</td>
<td>(-1.40)</td>
<td>(-1.37)</td>
<td>(-1.35)</td>
</tr>
<tr>
<td>IND</td>
<td>-0.460***</td>
<td>-0.537***</td>
<td>-0.523***</td>
<td>-0.526***</td>
</tr>
<tr>
<td></td>
<td>(-4.22)</td>
<td>(-5.01)</td>
<td>(-4.93)</td>
<td>(-4.96)</td>
</tr>
<tr>
<td>INN</td>
<td>0.056**</td>
<td>0.053**</td>
<td>0.051**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.32)</td>
<td>(2.17)</td>
<td>(1.97)</td>
<td></td>
</tr>
<tr>
<td>Nc</td>
<td></td>
<td>-0.002</td>
<td>-0.008</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-0.09)</td>
<td>(-0.46)</td>
<td></td>
</tr>
<tr>
<td>INN-Nc</td>
<td>-0.013*</td>
<td>-0.023***</td>
<td></td>
<td>(-2.58)</td>
</tr>
<tr>
<td></td>
<td>(-1.78)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RDL</td>
<td></td>
<td></td>
<td></td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.59)</td>
</tr>
<tr>
<td>INN-RDL</td>
<td></td>
<td></td>
<td></td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.09)</td>
</tr>
<tr>
<td>Nc-RDL</td>
<td></td>
<td></td>
<td></td>
<td>-0.008*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(-1.89)</td>
</tr>
<tr>
<td>INN-Nc-RDL</td>
<td></td>
<td></td>
<td></td>
<td>0.004***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(2.61)</td>
</tr>
<tr>
<td>cons</td>
<td>2.478***</td>
<td>2.566***</td>
<td>2.553***</td>
<td>2.537***</td>
</tr>
<tr>
<td></td>
<td>(23.24)</td>
<td>(24.07)</td>
<td>(23.87)</td>
<td>(23.78)</td>
</tr>
<tr>
<td>N</td>
<td>546</td>
<td>546</td>
<td>546</td>
<td>546</td>
</tr>
<tr>
<td>Waldy²</td>
<td>897.61</td>
<td>883.31</td>
<td>875.09</td>
<td>901.12</td>
</tr>
</tbody>
</table>

4.3 Robustness test

To verify the reliability of the above research results, the independent variable measurement method is changed in this paper and the number of applications for patents of the invention is used instead of the total number of patent applications to measure innovation capability. According to Table 4, the above research results are still valid.
### Table 4: Robustness test and analysis results

<table>
<thead>
<tr>
<th></th>
<th>Model (1)</th>
<th>Model (2)</th>
<th>Model (3)</th>
<th>Model (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-0.250</td>
<td>-0.236</td>
<td>-0.243</td>
<td>-0.227</td>
</tr>
<tr>
<td></td>
<td>(-1.31)</td>
<td>(-1.21)</td>
<td>(-1.27)</td>
<td>(-1.18)</td>
</tr>
<tr>
<td>Size</td>
<td>-0.286***</td>
<td>-0.337***</td>
<td>-0.347***</td>
<td>-0.343***</td>
</tr>
<tr>
<td></td>
<td>(-5.46)</td>
<td>(-5.76)</td>
<td>(-5.97)</td>
<td>(-5.88)</td>
</tr>
<tr>
<td>Debt</td>
<td>0.217</td>
<td>0.177</td>
<td>0.138</td>
<td>0.073</td>
</tr>
<tr>
<td></td>
<td>(0.90)</td>
<td>(0.72)</td>
<td>(0.56)</td>
<td>(0.29)</td>
</tr>
<tr>
<td>Inst</td>
<td>0.014***</td>
<td>0.014***</td>
<td>0.015***</td>
<td>0.015***</td>
</tr>
<tr>
<td></td>
<td>(2.00)</td>
<td>(2.05)</td>
<td>(2.19)</td>
<td>(2.19)</td>
</tr>
<tr>
<td>Indep</td>
<td>0.153</td>
<td>0.270</td>
<td>0.302</td>
<td>0.347</td>
</tr>
<tr>
<td></td>
<td>(0.25)</td>
<td>(0.44)</td>
<td>(0.49)</td>
<td>(0.56)</td>
</tr>
<tr>
<td>Stat</td>
<td>-0.062</td>
<td>-0.152</td>
<td>-0.233</td>
<td>-0.365</td>
</tr>
<tr>
<td></td>
<td>(-0.15)</td>
<td>(-0.38)</td>
<td>(-0.55)</td>
<td>(-0.86)</td>
</tr>
<tr>
<td>Dual</td>
<td>-0.126*</td>
<td>-0.120*</td>
<td>-0.118</td>
<td>-0.120</td>
</tr>
<tr>
<td></td>
<td>(-1.77)</td>
<td>(-1.68)</td>
<td>(-1.61)</td>
<td>(-1.61)</td>
</tr>
<tr>
<td>IND</td>
<td>-0.460***</td>
<td>-0.532***</td>
<td>-0.522***</td>
<td>-0.530***</td>
</tr>
<tr>
<td></td>
<td>(-4.22)</td>
<td>(-4.76)</td>
<td>(-4.72)</td>
<td>(-4.78)</td>
</tr>
<tr>
<td>INN</td>
<td>0.045*</td>
<td>0.043</td>
<td>0.041</td>
<td>0.041</td>
</tr>
<tr>
<td></td>
<td>(1.66)</td>
<td>(1.59)</td>
<td>(1.46)</td>
<td>(1.46)</td>
</tr>
<tr>
<td>Nc</td>
<td>0.005</td>
<td>-0.001</td>
<td>-0.015</td>
<td>-0.019**</td>
</tr>
<tr>
<td></td>
<td>(0.27)</td>
<td>(-0.07)</td>
<td>(-1.94)</td>
<td>(-2.15)</td>
</tr>
<tr>
<td>INN-Nc</td>
<td>-0.015*</td>
<td>-0.019**</td>
<td>-0.019</td>
<td>-0.019**</td>
</tr>
<tr>
<td></td>
<td>(1.66)</td>
<td>(1.59)</td>
<td>(1.46)</td>
<td>(1.46)</td>
</tr>
<tr>
<td>RDL</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INN-RDL</td>
<td>0.005</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.92)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nc-RDL</td>
<td>-0.007</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-1.46)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INN-Nc-RDL</td>
<td>0.003*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.82)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_cons</td>
<td>2.478***</td>
<td>2.556***</td>
<td>2.551***</td>
<td>2.539***</td>
</tr>
<tr>
<td></td>
<td>(23.24)</td>
<td>(23.00)</td>
<td>(22.85)</td>
<td>(22.61)</td>
</tr>
<tr>
<td>N</td>
<td>546</td>
<td>546</td>
<td>546</td>
<td>546</td>
</tr>
<tr>
<td>Waldχ²</td>
<td>897.61</td>
<td>873.76</td>
<td>870.92</td>
<td>875.58</td>
</tr>
</tbody>
</table>

### 5 Conclusions and Enlightenment

#### 5.1 Conclusions

This paper analyses and tests the impact of corporate innovation capability on corporate performance from the perspectives of value cognition and punctuated equilibrium. Also, it examines the moderating effect of internal cognitive characteristics of organizations on their innovation capability and corporate performance, as well as the moderating effect of the joint interaction between R&D leap and internal cognitive characteristics on innovation capability and corporate performance. The following three conclusions are finally drawn in this paper: Corporate innovation capability makes a positive impact on corporate performance; value cognition complexity negatively regulates the positive relationship between innovation capability and corporate performance and the relationship between innovation capability and corporate
performance is jointly regulated by R&D leap and value cognition complexity. In the case of a small R&D leap and a low value cognition complexity, this relationship is positive.

5.2 Theoretical Contribution

First, the two branches of the strategic management field, namely capability and cognition, are connected. The two major schools in the field of strategic management have been developing on parallel but separate paths for a long time, with a lack of connections and integration. In fact, the nature and usefulness of corporate capabilities are constrained by managers’ cognitive factors.

Second, research is made from the perspective of R&D leap in response to the research gap of “there has been no matching between R&D leap and value cognition characteristics to study its effect on corporate innovation capability improvement”. The allocation method of the value cognition complexity and R&D leap that helps improve corporate performance is found, which deepens value cognition research. Most of the previous value cognition research are based on the impact of internal resources of organizations on cognition but ignores the effect of the overall characteristics of cognition, as a group of cognitive mode, on corporate innovation.

6 SHORTCOMINGS AND FUTURE DIRECTION

First, due to the small number of AI listed companies from 2011 to 2021, there are only 73 enterprise samples, with small sample size. In the future, more enterprise samples can be studied. Second, some data are obtained by manually encoding the annual reports and social responsibility reports of these listed companies and they involve subjectivity and deviations to some extent.

REFERENCES


Research on the Relationship Between Absorptive Capacity and Innovation Capacity of AI enterprises: the Moderating Effect of R&D Leap and Value Cognition Complexity

Litian Chen1,a, Rui Wu1,b*, Yufeng Wang1,c, Junhong Zhong1,d
daisy0123@126.coma, *13349076657@163.comb, 493125335@qq.comc, 1579747581@qq.comd
School of Business Administration, Zhejiang Gongshang University, Hangzhou, Zhejiang, China1

Abstract: Artificial intelligence is the product of information technology at an advanced stage. As a key driver of the new round of global technological revolution and industrial transformation, artificial intelligence is building new momentum. For artificial intelligence enterprises, the key to maintain competitiveness in a changing environment lies in the improvement of Innovation capacity. To make effective innovations, they need to enhance their absorptive capability, acquire more knowledge from the outside world and break with tradition. From the perspective of punctuated equilibrium theory and value cognition, this paper, using the data of 84 listed companies of artificial intelligence concept stocks from 2011 to 2021 as a sample, analyses and examines the impact of absorptive capacity on Innovation capacity, and the moderating effect of R&D leap and value cognition complexity on the above relation. The results show that absorptive capacity positively affects Innovation capacity, that R&D leap negatively moderates the relationship between absorptive capacity and Innovation capacity, and that the joint interaction between value cognition complexity and R&D leap positively moderates the relationship between absorptive capacity and Innovation capacity. The findings of this paper provide guidance for enterprises to formulate innovation improvement strategies in a complex and changeable environment.

Keywords: Absorptive Capacity, Innovation Capacity, R&D Leap, Value Cognition.

1 INTRODUCTION

With the continuous improvement of computer hardware performance and computing technology, artificial intelligence technology and its applications are faced with significant opportunities [12]. As the field of artificial intelligence is experiencing rapid changes in an unpredictable environment, enterprises need to continuously absorb, acquire, and use external knowledge and resources and upgrade their technical foundations to survive and develop in the rat race. Meanwhile, the R&D leap resulting from the alternation and conversion of exploratory innovation and exploitative innovation-based on the punctuated equilibrium theory-in time helps achieve the discontinuous equilibrium of technological innovation and the effective management of R&D resources [10]. The ultimate goal is to maximize innovation output with high-value enterprise capabilities. However, enterprise innovation activities involve many value creation links [17]. Enterprises need to control the R&D leap through the value cognition system...
to effectively allocate resources and to improve the Innovation capacity of enterprises. In the era of digital intelligence, it is of vital importance for artificial intelligence enterprises to choose an appropriate R&D leap based on the value cognition framework to improve the synergy between the absorptive capacity and Innovation capacity. However, there are few research on this subject.

First, the relationship between absorptive capacity and Innovation capacity remains controversial. At present, there are different views in the academic community on the relationship between absorptive capacity and enterprise innovation. Most scholars believe that enterprise innovation requires the improvement of absorptive capacity of external knowledge, which can reduce innovation risks and improve innovation efficiency [1, 7]. However, some enterprises began to consider the negative impact of absorptive capacity in practice [11]. Instead of denying the role and value of absorptive capacity, they focus on the application scenarios of absorptive capacity.

Second, based on the punctuated equilibrium theory, enterprise innovation activities can be divided into two forms: exploratory innovation and exploitative innovation, and the leap in R&D investment is regarded as a sign of the conversion of the two, as well as the positive feedback that enterprises make according to its own demand and the market development. Previous studies mostly focused on the relationship between the jump in R&D investment and corporate performance [9], rather than on the relationship between it as a contextual factor and Innovation capacity. Accurately grasping the right R&D opportunity and rationally allocating R&D resources are the keys to continuously strengthen enterprise Innovation capacity. Therefore, it is necessary to take the above into consideration to study the boundary conditions of the relationship between the absorptive capability and Innovation capacity of artificial intelligence enterprises.

To fill the research gap, this paper uses the data of 84 listed companies with artificial intelligence concept stocks from 2011 to 2021 to study the relationship between the absorptive capacity and Innovation capacity of artificial intelligence companies, as well as the moderating effect of R&D leap and value cognition complexity on this relationship from the perspective of punctuated equilibrium theory and value cognition.

2 THEORETICAL BASIS AND HYPOTHESIS

2.1 Absorptive Capacity and Innovation Capacity

Amid increasing external uncertainties and turbulence, enterprises need to constantly upgrade their technological foundations. Absorptive capacity is essential for enterprises to acquire and apply external knowledge [15]. The absorptive capacity and Innovation capacity are closely related. The Innovation capacity of an enterprise is built on its own R&D activities [8] and external knowledge [4]. The application of external knowledge helps an enterprise strengthen its Innovation capacity, because they can improve the R&D efficiency [10] and success rate and optimize new products or new procedures. Therefore, the ability to absorb and utilize new external knowledge is crucial for enterprises to strengthen innovation and competitiveness. Enterprises with a high absorptive capacity usually actively identify and take advantage of market opportunities, effectively use external knowledge to stimulate innovations, and convert
it into innovation output, thereby promoting the Innovation capacity [22]. Based on this, the following assumption is made:

H1: Absorptive capacity positively affects the Innovation capacity of enterprises

2.2 The Moderating Effect of R&D Leap on the Relationship between Absorptive Capacity and Innovation Capacity

According to the punctuated equilibrium theory, the transition between exploitative innovation and exploratory innovation is regarded as a R&D leap, and such a leap in capability trajectory between exploitation and exploration is risky [2]. When an enterprise has a small R&D leap, they have made little breakthroughs in existing technology paths, which means they applied the external knowledge and resources they had acquired to existing innovation practices, and had more opportunities to improve Innovation capacity. When an enterprise has a large R&D leap, they have made significant breakthroughs in the existing paths, which means they constantly faced organizational inertia and core rigidity in process of innovation [18], experience and historical data didn’t work for new external exploration, they failed to reasonably apply external information and internal knowledge to innovation activities. Based on this, the following assumption is made:

H2: The R&D leap negatively adjusts the relationship between absorptive capacity and Innovation capacity, that is, the smaller the R&D leap, the easier to strengthen the positive relationship between absorptive capacity and Innovation capacity.

2.3 The Impact of the Joint Interaction of R&D Leap and Value Cognition Complexity on the Relationship between Absorptive Capacity and Innovation Capacity

The complexity of value cognition indicates the number of value-creating links considered by the enterprise in the process of decision-making cognition. The higher the complexity of value cognition, the more external stimulus enterprises need to consider, the more value links the managers need to focus on. In this context, if enterprises make significant R&D leap, they need to spend a lot of time to reallocate internal resources, so they fail to seize opportunities in time. Too many value links and too much resource allocation result in waste of resources, technology, and time, thus reducing the Innovation capacity. On the contrary, when the complexity of value cognition is low, managers need to consider less value links, so that enterprises can concentrate external knowledge and internal resources to seek suitable opportunities and directions. Therefore, small R&D leap can save time, reduce sunk costs and rationally allocate resources to enterprise innovation activities, thereby improving the Innovation capacity. Based on this, the following assumption is made.

H3: The joint interaction between R&D leap and value cognition complexity positively moderates the relationship between absorptive capacity and Innovation capacity. In the case of a small R&D leap and a low value cognition complexity, this relationship is positive.

The theoretical model of this paper is as follows:
3 METHODS AND MATERIALS

3.1 Samples and Data Sources

This paper uses the listed companies of the artificial intelligence concept stocks of Straight Flush from 2011 to 2021 as a sample. For more accurate measurement, we preliminarily excluded the data of companies marked with ST, ST*, PT and other signs of poor performance and of companies with missing or incomplete data. Finally, we obtained a sample of 84 companies and 756 specific observations.

This article uses stata 17.0 to process the data. In addition, 2% and 98% winsorize all continuous variables of the sample data to avoid the impact of outliers on the robustness of the results.

3.2 Variable Measurement

3.2.1 Dependent Variable: Innovation Capacity (IC)

Considering the innovation behavior of enterprises, patents serve as an important indicator of innovation quality. Innovation is to facilitate business operations of enterprises. In practice, the wider the application of a patent, the greater its value to the enterprise, and the more the enterprise is willing to pay for the patent. Therefore, patents are regarded as an indicator of enterprise quality. This paper draws on the research of Yang D et al. (2019) to measure the Innovation capacity of enterprises by the number of invention patent applications.

3.2.2 Independent Variable: Absorptive Capacity (AC)

The absorptive capacity of an enterprise is closely related to its R&D investment. Cohen and Levinthal (1990) proposed that the absorptive capacity of an enterprise can be measured by its own R&D investment, which revealed that Chinese enterprises have a relatively high conversion efficiency from R&D investment to innovation patents. Based on the research of Wu et al. (2016), this paper represents the absorptive capacity of enterprises by the intensity of R&D investment, which is expressed as the proportion of R&D investment in operating income.
3.2.3 Moderating Variable: R&D Leap (RDL)

This paper makes reference with the research of Mudambi and Swift (2014) and Swift (2016), and refers to their measurement methods of R&D leap. The specific steps are shown as follows:

Step 1: Calculate the autoregressive model residual \( u_{itn} \) of the ith enterprise in year t and then proceed to the modeling of the next step;

Step 2: Calculate the GARCH model residual \( e_{itn} \) of the ith enterprise in year t, which measures the degree to which the enterprise’s R&D expenditure in the year deviates from the predicted value that shows its historical trend;

Step 3: Then, calculate the studentized residual \( e_{itn}^{(stud)} \) for the R&D expenditure GARCH model of the ith enterprise in year t for subsequent comparative research. The specific calculating formula is shown as follows:

\[
e_{itn}^{(stud)} = \frac{e_{itn}}{s_{itn} \sqrt{1 - h_{int}}}
\]

where \( s_{it} \) is the standard deviation of \( e_{itn} \) and \( h_{int} \) is the impact of \( u_{itn} \) of the ith enterprise in year t on the entire estimation.

Step 4: Compare the absolute value of studentized residuals of each enterprise during the ten years from 2011 to 2020 and find the maximum value \( e_{itn}^{(max)} \) during the observation, which is the R&D leap of the ith enterprise.

3.2.4 Moderating Variable: Value Cognitive Complexity (Nc)

Based on the research of Nadkarni and Narayanan (2007), referring to the coding research design of Wu (2011), this paper uses content research methods to depict the value cognitive complexity of Chinese enterprises in the R&D process. The value creation activities of the enterprise are divided into two aspects: one is the main activities (Np), including R&D, design, supply, production, marketing and service; The second is support activities (Ns), covering information, relationship, operation and manpower. For each sample enterprise, we measure the total number of concepts created by the enterprise value chain.

Step 1: Identify the statements in the annual report and social responsibility report. That is, read the company’s annual report and corporate social responsibility report in the year of research and development, identify the statement segments related to the strategic innovation plan, and refer to the coding vocabulary summarized by Wu (2011) to record the statements of the annual report and social responsibility report on the main value creation and supporting factors in the innovation strategic plan segment.

Step 2: Calculate the complexity. After completing the above identification of the annual report and social responsibility report, this paper adds the number of main links and the number of support links to get the final total number of links according to the recorded phrases. This count is the value of cognitive complexity. The specific calculation formula is as follows:

\[ Nc = Np + Ns \]
If the complexity value calculated finally is larger, it means that the enterprise considers more value creation links, which proves that the cognitive complexity of the enterprise in the decision-making process is higher.

The specific vocabulary is shown in Table 1

3.2.5 Control Variable

This study selects control variables based on the literature of innovation research. Firstly, the strategic deployment of an enterprise will affect its innovation output. Therefore, this paper selects the following four control variables: the proportion of independent directors (Indep), the proportion of institutional ownership (INST), the proportion of state ownership (SOSP), and the dual role of the CEO (Dual). Secondly, since the innovation output of enterprises will be affected by enterprise resources, etc., this paper also regards the following as control variables: the age, the size, the asset-liability ratio (Lev), the bankruptcy distance (Zscore), and the growth. Finally, industry and year are taken as control variables, because they affect the innovation output of enterprises [3].

4 DATA RESULTS

4.1 Descriptive Statistics

This paper first performs descriptive analysis on the uncentered variables to obtain the average, standard deviation, etc., and then uses the Person correlation coefficient to describe the correlation between variables. It can be seen from Table 2 that the average values of Innovation capacity, absorptive capacity, R&D leap and cognitive value complexity are 2.712, 10.260, 2.201, 4.763 respectively. Additionally, it can be seen from Table 2 that the correlation coefficients of most variables are less than 0.8. Then, we examined the variance inflation factors of all variables, and found that all VIF values are less than 5, and the average value is 1.35, which indicates that the model in this paper will not incur serious multicollinearity problems.

4.2 Hypothesis Testing

To avoid problems of heteroscedasticity, serial correlation and cross-sectional correlation caused by the panel data used by this study, this paper uses Driscoll-Kraay (referred to as D-K) standard error with stata17.0 for estimation to achieve unbiased, consistent and effective results. As the results of Hausman test denied the null hypothesis, we used a fixed effect model. In addition, the interaction variables in the model were centralized to avoid multicollinearity.

Table 1: Coding vocabulary of value creation link

<table>
<thead>
<tr>
<th>Value creation links</th>
<th>Key words</th>
</tr>
</thead>
<tbody>
<tr>
<td>R&amp;D</td>
<td>Research, R&amp;D, scientific research, manufacture, development, etc.</td>
</tr>
<tr>
<td>Design</td>
<td>Design, planning, etc.</td>
</tr>
<tr>
<td>Production</td>
<td>Processing, OEM, smelting, rough refining, refining, fabrication, assembly, synthesis, production, etc.</td>
</tr>
<tr>
<td>Marketing</td>
<td>Publicity, development, expansion, promotion,</td>
</tr>
</tbody>
</table>
market, marketing, advertising, brand, image, underwriting, etc.
Exploration, mining, mining and beneficiation, procurement, purchase, transportation, logistics, freight, etc.
Supply
Market, repair, installation, debugging, representative, technical service, technical support, etc.
Service
Information collection, collection, research, investigation, understanding, opportunities, risks, analysis, feedback, briefing, reporting, mastery, etc.
Information management
Maintenance, repair, installation, debugging, representative, technical service, technical support, etc.
Relationship management
Contact, communication, coordination, relationship, organization, communication, cooperation, alliance, investment promotion, etc.
Operation management
Supervision, supervisor, control, finance, operation, project management, etc.
Support links
Raising, recruitment, talents, education, training, exchange, assignment, labor, personnel, etc.
Operation management

Table 2: Descriptive statistics of variables and Pearson correlation analysis

<table>
<thead>
<tr>
<th>Variable</th>
<th>IC</th>
<th>AC</th>
<th>RD</th>
<th>LNc</th>
<th>Size</th>
<th>Age</th>
<th>Indpe</th>
<th>INS</th>
<th>SOS</th>
<th>Gro</th>
<th>wth</th>
<th>Lev</th>
<th>Dua</th>
<th>Zsc</th>
</tr>
</thead>
<tbody>
<tr>
<td>IC</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AC</td>
<td>0.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RD</td>
<td>0.16</td>
<td>0.01</td>
<td>0.16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LNc</td>
<td></td>
<td></td>
<td></td>
<td>0.21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size</td>
<td>0.61</td>
<td>0.20</td>
<td>0.26</td>
<td></td>
<td>0.61</td>
<td>0.21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.14</td>
<td>0.13</td>
<td>0.06</td>
<td>0.14</td>
<td>0.14</td>
<td>0.07</td>
<td>0.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indpe</td>
<td>0.10</td>
<td>0.03</td>
<td>0.15</td>
<td>0.05</td>
<td>0.14</td>
<td>0.28</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOS</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
<td>0.05</td>
<td></td>
<td>0.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Growth</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lev</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dual</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zsc</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3: This caption has more than one line so it has to be set to justify.

<table>
<thead>
<tr>
<th>Model (1)</th>
<th>Model (2)</th>
<th>Model (3)</th>
<th>Model (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.003</td>
<td>0.155</td>
<td>0.209</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.70)</td>
<td>(0.95)</td>
</tr>
<tr>
<td>Size</td>
<td>0.831***</td>
<td>0.864***</td>
<td>0.862***</td>
</tr>
<tr>
<td></td>
<td>(17.23)</td>
<td>(18.62)</td>
<td>(18.55)</td>
</tr>
<tr>
<td>LEV</td>
<td>-0.332</td>
<td>0.084</td>
<td>0.189</td>
</tr>
<tr>
<td></td>
<td>(-1.11)</td>
<td>(0.28)</td>
<td>(0.64)</td>
</tr>
<tr>
<td>Growth</td>
<td>-0.460***</td>
<td>-0.343***</td>
<td>-0.347***</td>
</tr>
<tr>
<td></td>
<td>(-4.77)</td>
<td>(-3.53)</td>
<td>(-3.59)</td>
</tr>
<tr>
<td>INST</td>
<td>1.575**</td>
<td>1.555**</td>
<td>1.612**</td>
</tr>
<tr>
<td></td>
<td>(2.15)</td>
<td>(2.15)</td>
<td>(2.25)</td>
</tr>
<tr>
<td>Indep</td>
<td>-0.936</td>
<td>-0.725</td>
<td>-0.793</td>
</tr>
<tr>
<td></td>
<td>(-1.42)</td>
<td>(-1.14)</td>
<td>(-1.26)</td>
</tr>
<tr>
<td>SOSP</td>
<td>-0.292</td>
<td>-0.299</td>
<td>-0.307</td>
</tr>
<tr>
<td></td>
<td>(-0.49)</td>
<td>(-0.52)</td>
<td>(-0.53)</td>
</tr>
<tr>
<td>Dual</td>
<td>-0.109</td>
<td>-0.102</td>
<td>-0.122</td>
</tr>
<tr>
<td></td>
<td>(-1.31)</td>
<td>(-1.26)</td>
<td>(-1.49)</td>
</tr>
<tr>
<td>Zscore</td>
<td>-0.001</td>
<td>-0.004</td>
<td>-0.006</td>
</tr>
<tr>
<td></td>
<td>(-0.17)</td>
<td>(-0.78)</td>
<td>(-1.03)</td>
</tr>
<tr>
<td>AC</td>
<td>0.037***</td>
<td>0.038***</td>
<td>0.036***</td>
</tr>
<tr>
<td></td>
<td>(5.92)</td>
<td>(6.41)</td>
<td>(6.22)</td>
</tr>
<tr>
<td>RDL</td>
<td>-0.030**</td>
<td>-0.022</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-1.99)</td>
<td>(-1.45)</td>
<td></td>
</tr>
<tr>
<td>AC_RDL</td>
<td>-0.004**</td>
<td>-0.004**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-2.49)</td>
<td>(-2.31)</td>
<td></td>
</tr>
</tbody>
</table>
4.2.1 Testing of the Relationship between Absorptive Capacity and Innovation Capacity

It can be seen from the model (2) in Table 3 that the regression coefficient between the absorptive capacity and the Innovation capacity is 0.037, P<0.01. Thus, hypothesis 1 has been verified, that is, the absorptive capacity of the enterprise has a significant positive impact on the Innovation capacity. The stronger the enterprise absorptive capacity, the more effective it can convert external knowledge into internal capital, thus promoting its Innovation capacity.

4.2.2 The Moderating Effect Of R&D Leap on the Relationship between Absorptive Capacity and Innovation Capacity

As seen from the model (3) in Table 3, the regression coefficient the product term of the absorptive capacity and R&D leap is -0.004, P<0.05. Thus, hypothesis 3 is verified. This shows that the R&D leap has a negative moderating effect on the relationship between absorptive capacity and innovation capacity, that is, when the R&D leap is low, the continuous improvement of absorptive capacity of the enterprise is conducive to the growth of its innovation capacity.

4.2.3 The Impact of the Joint Interaction of R&D Leap and Value Cognition Complexity on Enterprise Absorptive Capacity and Innovation Capacity

As seen from the model (4) in Table 3, the regression coefficient of the cubic product term of enterprise absorptive capacity, R&D leap, and value cognition complexity is 0.001, P<0.1, indicating that the R&D leap and value cognition complexity will jointly affect the relationship between absorptive capacity and innovation capacity. In addition, as seen from Table 3, when the R&D leap is small and the complexity of value cognition is low, the stronger the absorptive capacity of the enterprise, the more it can improve its innovation capacity. That is, in the case of low R&D leap and value cognition complexity, the enterprise absorptive capacity and innovation capacity have a stronger positive relationship. Thus, hypothesis 3 is verified.

4.3 Robustness Verification

In order to verify the reliability of the above results, this paper replaces the data with a lag of one period with the data of Innovation capacity with a lag of two periods. After a regression
analysis of the model, we obtained Table 4 with consistent results, which indicates that the results have a strong robustness.

Table 4 Robustness test and analysis results

<table>
<thead>
<tr>
<th></th>
<th>Model (1)</th>
<th>Model (2)</th>
<th>Model (3)</th>
<th>Model (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-0.156</td>
<td>0.044</td>
<td>0.113</td>
<td>0.129</td>
</tr>
<tr>
<td></td>
<td>(-0.63)</td>
<td>(0.18)</td>
<td>(0.46)</td>
<td>(0.53)</td>
</tr>
<tr>
<td>Size</td>
<td>0.733***</td>
<td>0.776***</td>
<td>0.776***</td>
<td>0.786***</td>
</tr>
<tr>
<td></td>
<td>(14.96)</td>
<td>(16.52)</td>
<td>(16.43)</td>
<td>(16.50)</td>
</tr>
<tr>
<td>LEV</td>
<td>0.753**</td>
<td>1.047***</td>
<td>1.123***</td>
<td>1.169***</td>
</tr>
<tr>
<td></td>
<td>(2.30)</td>
<td>(3.28)</td>
<td>(3.53)</td>
<td>(3.63)</td>
</tr>
<tr>
<td>Growth</td>
<td>-0.101</td>
<td>0.006</td>
<td>0.016</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td>(-0.97)</td>
<td>(0.05)</td>
<td>(0.16)</td>
<td>(0.15)</td>
</tr>
<tr>
<td>INST</td>
<td>1.684*</td>
<td>1.728**</td>
<td>1.927**</td>
<td>2.021**</td>
</tr>
<tr>
<td></td>
<td>(1.93)</td>
<td>(2.01)</td>
<td>(2.28)</td>
<td>(2.38)</td>
</tr>
<tr>
<td>Indep</td>
<td>0.185</td>
<td>0.459</td>
<td>0.371</td>
<td>0.366</td>
</tr>
<tr>
<td></td>
<td>(0.27)</td>
<td>(0.68)</td>
<td>(0.56)</td>
<td>(0.54)</td>
</tr>
<tr>
<td>SOSP</td>
<td>-0.226</td>
<td>-0.191</td>
<td>-0.138</td>
<td>-0.398</td>
</tr>
<tr>
<td></td>
<td>(-0.27)</td>
<td>(-0.23)</td>
<td>(-0.16)</td>
<td>(-0.48)</td>
</tr>
<tr>
<td>Dual</td>
<td>-0.178**</td>
<td>-0.175**</td>
<td>-0.199**</td>
<td>-0.185**</td>
</tr>
<tr>
<td></td>
<td>(-1.96)</td>
<td>(-1.98)</td>
<td>(-2.24)</td>
<td>(-2.06)</td>
</tr>
<tr>
<td>Z_score</td>
<td>0.008</td>
<td>0.005</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>(1.39)</td>
<td>(0.83)</td>
<td>(0.44)</td>
<td>(0.42)</td>
</tr>
<tr>
<td>AC</td>
<td>0.038***</td>
<td>0.040***</td>
<td>0.040***</td>
<td>0.040***</td>
</tr>
<tr>
<td></td>
<td>(5.38)</td>
<td>(5.96)</td>
<td>(5.90)</td>
<td></td>
</tr>
<tr>
<td>RDL</td>
<td>-0.037**</td>
<td>-0.034**</td>
<td>-0.034**</td>
<td>-0.044**</td>
</tr>
<tr>
<td></td>
<td>(-2.35)</td>
<td>(-2.11)</td>
<td></td>
<td>(-2.49)</td>
</tr>
<tr>
<td>AC_RDL</td>
<td>-0.004**</td>
<td>-0.004**</td>
<td>-0.004**</td>
<td>-0.004**</td>
</tr>
<tr>
<td></td>
<td>(-2.49)</td>
<td>(-2.57)</td>
<td></td>
<td>(-2.49)</td>
</tr>
<tr>
<td>Nc</td>
<td></td>
<td></td>
<td></td>
<td>-0.010</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(-0.45)</td>
</tr>
<tr>
<td>AC_Nc</td>
<td>0.003</td>
<td></td>
<td></td>
<td>(1.10)</td>
</tr>
<tr>
<td>RDL_Nc</td>
<td>0.010*</td>
<td></td>
<td></td>
<td>(1.76)</td>
</tr>
<tr>
<td>AC_RDL</td>
<td>0.002**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(2.17)</td>
</tr>
<tr>
<td>cons</td>
<td>2.617***</td>
<td>2.468***</td>
<td>2.510***</td>
<td>1.600***</td>
</tr>
<tr>
<td></td>
<td>(19.20)</td>
<td>(17.90)</td>
<td>(18.23)</td>
<td>(11.38)</td>
</tr>
<tr>
<td>year</td>
<td>Controlled</td>
<td>Controlled</td>
<td>Controlled</td>
<td>Controlled</td>
</tr>
<tr>
<td>IND</td>
<td>Controlled</td>
<td>Controlled</td>
<td>Controlled</td>
<td>Controlled</td>
</tr>
<tr>
<td>N</td>
<td>672</td>
<td>672</td>
<td>672</td>
<td>672</td>
</tr>
<tr>
<td>Wald χ²</td>
<td>595.83***</td>
<td>711.80***</td>
<td>751.88***</td>
<td>775.72***</td>
</tr>
</tbody>
</table>

5 CONCLUSION AND IMPLICATIONS

5.1 Conclusion and Discussion

From the perspective of punctuated equilibrium theory and value cognition, this paper analyses and examines the impact of enterprise absorptive capacity on Innovation capacity, and analyses...
how the R&D leap moderates the relationship between absorptive capacity and Innovation capacity, as well as how the joint interaction of R&D leap and value cognition complexity regulates the above relationship. This paper comes into three conclusions. Firstly, the absorptive capacity of enterprises positively affects the Innovation capacity. Strong absorptive capacity means that the enterprise can adapt to changes in the innovation environment and technology more quickly, and can digest, integrate and utilize internal and external resources more effectively, so as to improve its adaptability, Innovation capacity, and competitiveness. Secondly, the R&D leap negatively moderates the relationship between absorptive capacity and Innovation capacity. Large R&D leap may be detrimental to the management and allocation of R&D resources, thus reducing the resources required for the growth of Innovation capacity. Thirdly, the joint interaction of value cognition complexity and R&D leap positively moderates the relationship between absorptive capacity and Innovation capacity.

5.2 Theoretical Contributions

Firstly, this research discusses for the first time the relationship between absorptive capacity and Innovation capacity using R&D leap amplitude as a situational variable, thus deepening research on punctuated equilibrium theory. Previous studies have shown that companies with large R&D leap need to coordinate and allocate resources, resulting in greater waste. According to the discontinuous balance theory, exploratory innovation and exploitative innovation alternate in time to achieve a balance [10]. Based on the typical environment of artificial intelligence enterprises, this paper expands the application scenarios of discontinuous equilibrium, sheds new light on the operating mechanism between the three, and proposes a new perspective for further study of the relationship between absorptive capacity and Innovation capacity.

Secondly, in response to the lack of research on the impact of value cognition complexity on the relationship between absorptive capacity and Innovation capacity, this study differs from previous research [6], by innovatively regarding value cognition complexity as a cognitive model under the innovation framework for further research and analysis. It is found that the joint interaction of value cognition complexity and R&D leap has a certain impact on the relationship between absorptive capacity and Innovation capacity. This paper systematically studies the moderating effect of contextual factors on the relationship between absorptive capacity and Innovation capacity. By analysing these moderating effects, this study, to a certain extent, explains the reasons for the divergence of existing research on the relationship between absorptive capacity and Innovation capacity, providing guidance for future research.

5.3 Limitations and Future Directions

This study has the following limitations: firstly, part of the data were obtained by hand-coding the annual reports and social responsibility reports of listed companies, which, to a certain extent, are subjective and biased; secondly, this study only measures the two variables of absorptive capacity and Innovation capacity from one perspective. In the future, the indicators can be considered from a multi-dimensional perspective, and the research can be deepened from more perspectives.
Acknowledgment: This research was substantially supported by research grants from the Zhejiang philosophy and social science planning project (20NDJC099YB), National Natural Science Foundation of China (71972170).

REFERENCES

The Management Rights and Accounting Conservatism -- Data Mining and Data Analysis of 4404 Listed Companies

Chunyu Meng
1248115109@qq.com
Business school of Sichuan University Department of Accounting and Corporate Finance, China

Abstract: With the deep integration of Internet and economy, big data mining has brought more convenience to accounting practice and accounting related academic research. Stable accounting information is the basic condition for listed companies to carry out activities in accordance with the real internal situation, and the management is the most direct and significant external force in the influencing factors of the company's accounting conservatism. In order to explore the relationship between the two, this paper used Python to climb the annual report data of A-share listed companies to construct the variable of management power. Taking Chinese A-share listed companies from 2017 to 2020 as samples, this paper studied the influence of management power of listed companies on accounting conservatism. The empirical study showed that listed companies with more management power have worse performance of accounting conservatism index. The research conclusion of this paper has certain enlightenment significance for listed companies to improve accounting conservatism.

Keywords: Component, Management Rights, Accounting Conservatism, Big Data Analytics.

1 INTRODUCTION

Accounting information, as the indicator of enterprise development, is extremely important for enterprises, the whole capital market and even the social resource market. Robust accounting information can reduce information asymmetry and improve investment efficiency. As the managers of listed companies, accounting information is one of the important criteria for evaluating their business results. Therefore, the management, shareholders and other economic stakeholders all pay close attention to accounting information, which is provided by the management and is bound to be directly affected by the management. Therefore, in order to improve the overall information quality level of the capital market and safeguard the interests of all stakeholders, it is necessary to analyze the role direction of the management on accounting conservatism, and conduct empirical analysis with the data of listed companies to explore the correlation between the two.
Therefore, based on the realistic background of our country, this paper studied the relationship between management power and accounting conservatism on the basis of relevant scholars, the contribution of this article is: to improve the research system of influencing factors of accounting conservatism. At present, there are many studies on management power and accounting conservatism by relevant scholars, but no conclusion has been reached. The research content of this paper provides a broader idea for the development of related research on accounting conservatism, which is helpful to improve the research system of influencing factors of accounting conservatism. This paper constructs the relevant model, synthesizes the management power with the indicators of the five aspects that affect the management power, selects appropriate samples to carry out empirical research, and provides guidance for listed companies to improve the accounting conservatism from the perspective of the management power allocation based on the empirical conclusion.

2 LITERATURE REVIEW AND HYPOTHESIS PRESENTED

In his research, Roll(1986) believed that the management is usually overconfident and will overestimate the accuracy and comprehensiveness of the information they own, their personal management ability and the future business conditions of the enterprise, and then make irrational investment behaviors, which will bring losses to the development of the enterprise [1]. After studying media reports, managerial control rights and accounting conservatism of listed companies, Besley et al. (2002) concluded that media attention and related reports on listed companies with high managerial power concentration were significantly higher than those with low managerial power concentration. The restriction of public opinion on management is conducive to ensuring accounting conservatism [2]. Zygidopoulos et al. (2012) believe in their study that excessive concentration of power gives the management the ability to construct a compensation system conducive to its own interests, and at the same time, it will also help the management to implement earnings management, thereby damaging accounting conservatism [3]. Ahmed et al. (2012) found in their study that accounting conservatism would not be affected if the chairman and general manager of an enterprise were the same person, but this study could not prove that management power was unrelated to accounting conservatism, because the combination of the two positions was only one aspect of management power [4]. Cheng and Warfield(2015) pointed out in their paper that the ability of the board of directors to check and balance the general manager is inversely proportional to the concentration of the general manager's power. The more concentrated the general manager's power is, the more likely the company is to have internal transactions with its affiliated companies and reduce the accounting conservatism [5].

Zhao Yali and Gao Chengjing (2014) found that the management of listed companies would have an impact on the quality of accounting information. The stronger the management's ability to control the listed company's operation and management, investment decisions, compensation setting and the establishment of reward and punishment system, the lower the accounting conservatism [6]. According to the research of Lin Jing and Guo Junhang (2020), an important reason for the overconfidence of the management is their excessive power, which will ultimately greatly reduce the conservatism of the accounting policies of listed companies [7]. According to Luo Min (2015), corporate supervision loopholes provide necessary conditions...
for the management's rent-seeking power, and cases that eventually damage the overall interests of listed companies are common. The level of management's power determines the degree of rent-seeking power. Management means of power rent-seeking is usually human tampering with company profits, forge the good illusion, blinded by the supervisor of judgment, achieve the purpose of to raise their salary income, moreover may also directly formulate conducive to its own system, and by taking advantage of their office to unsound accounting policies, even random change of accounting policy, finally lead to accounting information disclosure sustainability is poor, Reduce the level of information robustness [8].

Management power as emerging theory at home and abroad, less research about the relation between management power and accounting conservatism, part of the study results show that the management of power have a negative effect on accounting conservatism, but there are still some scholars under the different perspective, it is concluded that the management power positive than the relationship between accounting conservatism and conclusion, Therefore, it is necessary to study them further.

The internal and external information asymmetry derived from the principal-agent relationship gives the management information advantage, provides the conditions for the management to use their power to control the quality of accounting information, and further provides convenient conditions for the management to seek personal interests. In addition, the excessive expansion of management power also provides conditions for management to reduce accounting conservatism. Therefore, the hypothesis is proposed as follows:

H1: The power of management is negatively correlated with accounting conservatism, that is, the greater the power of management, the lower the accounting conservatism of listed companies.

3 RESEARCH DESIGN

3.1 Data

The independent variable data of the empirical part of this paper comes from the annual reports of A-share listed companies that are crawled by Python software. The accounting conservatism of the dependent variable is downloaded from the CSMAR database. Excel2017 and Stata15.1 are used for sorting and empirical testing, and financial listed companies are excluded. After the above data processing, 4404 listed companies with a total of 11,349 samples from 2017 to 2020 were obtained for empirical research.

3.2 Variable Definitions

3.2.1 Accounting Conservatism

Because the paper mainly investigates the influence of management power on accounting conservatism, thus accounting conservatism is taken as the explained variable. Scholars at home and abroad have studied a variety of measurement methods of accounting conservatism. At present, scholars at home and abroad use Basu reverse regression model to measure
accounting conservatism. However, the application premise of Basu reverse regression model is that the capital market is completely effective, the development time of Chinese capital market is short, and it does not meet the basic application condition, so the model is not completely applicable to measure the accounting conservatism. Secondly, the accounting conservatism measured by Basu reverse regression model can only reflect the existence of accounting conservatism of listed companies, but cannot reflect the specific level of accounting conservatism. Therefore, this paper refers to the listed company robustness index calculated by ACF model.

<table>
<thead>
<tr>
<th>indicators</th>
<th>Relationship with management power</th>
<th>The values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duality</td>
<td>When the management also serves as the chairman, the management has more power</td>
<td>1 if two jobs are in one, 0 otherwise</td>
</tr>
<tr>
<td>Bdnum</td>
<td>The larger the scale of the board of directors, the weaker the supervision and the greater the power of the management</td>
<td>If it's greater than the sample mean, it's 1, otherwise it's 0</td>
</tr>
<tr>
<td>Proportion of independent directors</td>
<td>The larger the proportion of independent directors in the board of directors, the less power the management has</td>
<td>1 less than the sample mean, 0 otherwise</td>
</tr>
<tr>
<td>Management shareholding ratio</td>
<td>The management rights increase with the increase of the shares held by the listed company</td>
<td>If it's greater than the sample mean, it's 1, otherwise it's 0</td>
</tr>
<tr>
<td>Balance of ownership</td>
<td>The more dispersed the ownership structure, the more power the management has</td>
<td>When the ratio of the sum of shares held by the second to fifth largest shareholders to the shares of the largest shareholder is greater than 1, the value is 1; otherwise, the value is 0</td>
</tr>
</tbody>
</table>

### 3.2.2 Management Rights

For the measurement of management Power, refer to the research of Li Yingying (2019) [9] and other scholars, and the final paper simply adds the five indicators in the table below to synthesize the comprehensive indicator power of management power.

### 3.2.3 Control Variables

Referring to the research literature on accounting conservatism by domestic and foreign scholars such as Liu Xinmei (2016) and Jia Wanjun et al. (2017), the natural logarithm of total assets, asset-liability ratio, mark-to-book ratio and return on equity are taken as control variables [10-11]. According to the research of relevant scholars, the natural logarithm of total assets, asset-liability ratio and market-to-book ratio are selected as control variables. The variables are selected as shown in the following table.
## TABLE II. VARIABLES

<table>
<thead>
<tr>
<th>Variable types</th>
<th>Variable symbol</th>
<th>Variable name</th>
<th>Variable definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>The dependent variable</td>
<td>ACF</td>
<td>Accounting conservatism</td>
<td>Accounting conservatism as measured by ACF model</td>
</tr>
<tr>
<td>The independent variables</td>
<td>Power</td>
<td>Right of management</td>
<td>Comprehensive indicators of management power</td>
</tr>
<tr>
<td>Control variables</td>
<td>Size</td>
<td>Size of listed company</td>
<td>Log of total assets at year-end</td>
</tr>
<tr>
<td></td>
<td>Roa</td>
<td>Net profit margin on total assets</td>
<td>Net profit/total asset balance</td>
</tr>
<tr>
<td></td>
<td>Roe</td>
<td>Return on equity</td>
<td>Net profit/shareholders’ equity balance</td>
</tr>
<tr>
<td></td>
<td>Lev</td>
<td>Asset liability ratio</td>
<td>Total assets/total liabilities</td>
</tr>
<tr>
<td></td>
<td>Mtb</td>
<td>Market to book ratio</td>
<td>Ending market value to book value of total assets</td>
</tr>
</tbody>
</table>

### 3.3 Model

According to hypothesis 1, regression model 1 was constructed with management (Power) as explanatory variable.

\[
ACF_{ij} = \beta_0 + \beta_1 \text{Power}_{ij} + \beta_2 \text{Size}_{ij} + \\
\beta_3 \text{Roa}_{ij} + \beta_4 \text{Roe}_{ij} + \beta_5 \text{Lev}_{ij} + \beta_6 \text{Mtb}_{ij} + \epsilon_{ij}
\]

As for the model 1 concerned, if \( \beta_1 \) significantly less than 0, it indicates that the power of management of listed company is bigger, the accounting conservatism of listed company is lower.

### 4 THE EMPIRICAL RESULTS

#### 4.1 Descriptive Statistics

As shown in the following table, the mean value of accounting conservatism is -0.558, the minimum value is -3.592, and the maximum value is 3.946, indicating that there are great differences in accounting conservatism among listed companies, and the overall accounting conservatism level needs to be improved. The mean value of management rights is 2.187, the maximum value is 5, and the minimum value is 0, indicating that the overall fluctuation of management rights of listed companies is large. The mean value of company size is 22.371, the maximum value is 26.324, and the minimum value is 19.923, indicating that the difference of company size is small. The mean value of net profit margin of total assets is 0.034, and the difference between the minimum value and the maximum value is large, indicating that the utilization effect of total assets of Chinese listed companies is quite different. The general fluctuation of return on equity is relatively large; The mean value of the asset-liability ratio is 0.42, the maximum value is 0.994, and the minimum value is 0.008, indicating that the overall solvency of Chinese listed companies is good. Market to book ratio overall difference is large.
### TABLE III. DESCRIPTIVE STATISTICS

<table>
<thead>
<tr>
<th>Variable</th>
<th>Observations</th>
<th>Mean</th>
<th>Std. dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACF</td>
<td>11349</td>
<td>-0.558</td>
<td>0.990</td>
<td>-3.592</td>
<td>3.946</td>
</tr>
<tr>
<td>Power</td>
<td>11349</td>
<td>2.187</td>
<td>1.177</td>
<td>0.000</td>
<td>5</td>
</tr>
<tr>
<td>Size</td>
<td>11349</td>
<td>22.371</td>
<td>1.304</td>
<td>19.923</td>
<td>26.324</td>
</tr>
<tr>
<td>Roa</td>
<td>11349</td>
<td>0.034</td>
<td>0.094</td>
<td>-2.646</td>
<td>0.786</td>
</tr>
<tr>
<td>Roe</td>
<td>11349</td>
<td>0.016</td>
<td>1.010</td>
<td>-53.038</td>
<td>2.324</td>
</tr>
<tr>
<td>Lev</td>
<td>11349</td>
<td>0.420</td>
<td>1.196</td>
<td>0.008</td>
<td>0.994</td>
</tr>
<tr>
<td>Mtb</td>
<td>11349</td>
<td>1.796</td>
<td>1.855</td>
<td>0.038</td>
<td>33.232</td>
</tr>
</tbody>
</table>

### 4.2 Linear Regression Analysis

The results in the following table show that the partial regression coefficient between management Power (Power) and corporate accounting conservatism (ACF) is -0.014, which is significant at the 10% level, indicating that there is a significant negative correlation between them: the greater the management power, the higher the accounting conservatism of listed companies. Therefore, in the regression model, the correlation coefficient and significance level between management power and accounting conservatism are the main concerns, so hypothesis 1 passes the test.

### TABLE IV. LINEAR REGRESSION ANALYSIS

<table>
<thead>
<tr>
<th>Variable</th>
<th>ACF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Power</td>
<td>-0.014*</td>
</tr>
<tr>
<td></td>
<td>(-1.79)</td>
</tr>
<tr>
<td>Size</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>(0.58)</td>
</tr>
<tr>
<td>ROA</td>
<td>0.565***</td>
</tr>
<tr>
<td></td>
<td>(4.46)</td>
</tr>
<tr>
<td>ROE</td>
<td>-0.006</td>
</tr>
<tr>
<td></td>
<td>(-0.60)</td>
</tr>
<tr>
<td>Lev</td>
<td>0.100*</td>
</tr>
<tr>
<td></td>
<td>(1.65)</td>
</tr>
<tr>
<td>Mtb</td>
<td>-0.035***</td>
</tr>
<tr>
<td></td>
<td>(-5.39)</td>
</tr>
<tr>
<td>Contant</td>
<td>-0.639***</td>
</tr>
<tr>
<td></td>
<td>(-3.35)</td>
</tr>
</tbody>
</table>

\* t statistics in parentheses

* p < 0.1, ** p < 0.05, *** p < 0.01

### 5 CONCLUSIONS

Through the empirical study, the following conclusions are drawn: accounting conservatism decreases with the increase of management power in listed companies. As the principal, the management is also an individual with self-interest motive, so it is inevitable to play favoritism
and improve their own compensation by manipulating the company's income cost and other related information, thus causing adverse impact on the accounting conservatism of the company. The more power the management has, the more influence it has on enterprise operation and management, and the more likely it is to manipulate accounting information and reduce accounting conservatism.

Shortcomings of the paper: Management power is influenced by subjective and objective factors, not limited to the indicators selected in the paper. Factors such as the process of marketization, the level of knowledge and moral accomplishment of the management, the supervision of the media and the masses, and the institutional environment also have an impact on the power of the management. Therefore, the indicators selected in this paper may not be sufficient to comprehensively measure the power of the management. The results would be more convincing if they were further explored in subsequent studies taking other factors into account.

REFERENCES

Statistical Studies on the Influencing Factors of College Students' Mystery Box Consumption Based on Path Analysis: A Case Study of Guangzhou

Hua He¹, a, Wangling Feng¹, b
lotus@stu.scau.edu.cn¹, 201823180403@stu.scau.edu.cn²

College of Economics and management, South China Agricultural University, No. 483, Wushan Road, Tianhe District, Guangzhou, Guangdong, China¹

Abstract: Based on the theory of behavioral economics, this paper explores the influencing factors of mystery boxes, a new popular commodity that college students are keen on buying. In this paper, PLS-PM model, Apriori correlation analysis and logistic regression are adopted to further explore the psychological perceived value that can approximately represent the mystery boxes sold by college students. The research shows that college students' mystery boxes consumers pursue personal satisfaction and rarely blindly follow the trend of consumption. They care most about the "surprise" brought by opening mystery boxes. At the same time, the research results show that for mystery boxes, increasing the social popularity, enhancing their IP cultural attributes, and increasing the variety of rare or special products can effectively enhance consumers' purchase intention to mystery boxes. Finally, since the consumption mental account is set for mystery boxes in advance, college students will unconsciously consume the mystery boxes irrationally.

Keywords: Mystery Boxes, Behavioral Economics, Factor Analysis, PLS-PM, Association Analysis, Logistic Regression.

1 INTRODUCTION

Mystery boxes, a new type of goods equipped with IP culture, have become popular among young consumers in recent years. Mystery boxes² is becoming a common marketing method, with companies in different fields launching mystery boxes. In 2016, Pop-Mart, whose main business is fashionable toy IP incubation operation, began to introduce mystery boxes related products into China. According to the fashionable toy industry research report of a number of securities companies, the market size of fashionable toy has reached 80 billion in 2022, of which mystery boxes represented by Pop-Mart account for 40% of it. However, it is important to note that mystery boxes economy suffers from false advertising, poor quality products, etc. In addition, mystery boxes hide limited editions in marketing, which is very easy for speculators to exploit loopholes for windfall profits, and the price of products is several times the original price, causing market chaos. College students, as a group that pursues fashion and is prone to impulsive consumption, lead to a challenging topic at the moment when mystery boxes are hot: what exactly attracts people to consume and buy again behind the mystery boxes economy.
2 LITERATURE REVIEW

Fashionable toy is an adult toy that integrates the multi-element concepts of art, design, trend, painting, sculpture and so on. Mystery Boxes is an innovative fashionable toy that is popular for carrying a rich and diverse trend culture.

In the new era of consumption, consumers pay more attention to the emotional support brought by products, and mystery boxes have the unique "social and fashionable" properties of toys, which will meet the needs of consumers in the new era. At the same time, in the process of purchasing and opening mystery boxes, consumers get a pleasant feeling, which makes the whole purchase process more interesting [1].

However, the "hidden limit" mechanism in mystery boxes is very easy to lead to the confusion of "mystery boxes", which can be easily used by speculators [2]. For consumers, in real life, it is inevitable to be influenced by intuition, situation, habits, customs and other factors, and make some irrational economic behaviors that violate the economic operation law [3].

3 THE ESTABLISHMENT OF RESEARCH METHODS AND MODELS

3.1 Research Design

3.1.1 Research Content

This paper studies mystery boxes economy with perceived value. Perceived value refers to the overall evaluation of a product or service by customers after comparing the perceived utility of the product or service with the sacrifice they have made in obtaining the product or service [4]. From the perspective of mystery boxes economy, perceived value is the value given to mystery boxes by consumers who experience them from mystery boxes. It is directly reflected in the overall satisfaction, cost performance and psychological gap of consumers on mystery boxes, and these three points directly reflect the psychological feelings of consumers after purchase.

Based on the Perceived Value of mystery boxes consumers, This paper proposes independent consumption, group psychology, symbolic consumption, Fashion Consumption, Operate conditioning:

Group psychology refers to the psychology that when people make decisions, their choices are easily influenced by the behavior and thinking styles of other decision makers, so as to weaken the significance of such decisions for individuals. Most college students are far away from their parents, have a growing sense of self, and have a certain income, so they are more independent in the process of consumption. Symbolic consumption refers to consumers not only consume the goods themselves, but also some cultural significance symbolized by the consumption of these goods, including the mood, aesthetic feeling, atmosphere, and emotion obtained by consumption. Fashion Consumption refers to the psychological experience that people pursue the forefront of the trend in the process of consumption, take the initiative to understand product information at the first time or pursue the latest products, hoping to have new and different and leading the trend. The Operate conditioning system refers to the...
dependence of consumers on mystery boxes, which is similar to that of tobacco and alcohol. The uncertain incentive of mystery boxes positively affects the willingness of consumers to continue to purchase.

3.1.2 Research Scheme

This paper first explored the influence of each factor on perceived value, analyzed and determined the influencing factors, and established the PLS-PM model to explain the correlation of factors according to the results. Then, the five association rules with the highest confidence were mined from the 15 variables through Apriori association analysis, and the eight factors involved in the five association rules were regarded as independent variables for logistic regression with purchase behavior, so as to find the influence degree of each variable on the dependent variable. Then, 15 variables were directly trained by random forest dichotomies to verify the logistic regression results. Finally, suggestions are given from three perspectives.

3.2 Sample Selection and Data Sources

3.2.1 Sampling Survey

In this paper, a self-designed questionnaire is used to collect data based on the theory of perceived value and the actual situation of mystery boxes consumption. The questionnaire is divided into six parts: personal information, mystery boxes consumption, consumption intention, post-purchase psychology, subjective judgment of mystery boxes and factors affecting consumers' "perceived value".

A random survey was conducted among students from different universities and grades in Guangzhou. The minimum sample size was determined to be 384 after pre-survey.

The official survey was sent out online. After screening, 435 questionnaire samples were collected, and the proportion of valid samples in the questionnaire was about 92.55%, which was in line with the expectation of the minimum sample size.

When exploring the perceived value of participating in the mystery boxes economy and its influencing factors, the data analysis object will be narrowed down to 251 samples of mystery boxes consumers. In the remaining chapters, a complete set of 427 valid questionnaires will be selected. The data samples filled by the respondents will be included in the analysis whether buying mystery boxes or not, so as to explore the characteristics of consumer groups.

3.2.2 Quality Control

Reliability and validity tests on questionnaire scales found that all Cronbach's alpha coefficients exceeded the minimum acceptable level of 0.7 and were internally consistent. The Cronbach's alpha coefficient value of the total scale is 0.894, greater than 0.7, and the questionnaire reliability is good. The KMO statistical analysis value is 0.832, and the spherical test result is significant (p<0.001), which is suitable for factor analysis.

3.3 PLS-PM Model Establishment

This paper explores the influence of various factors on the perceived value of the research group and establishes the PLS-PM model as shown in Figure 1.
The results of principal component analysis showed that the cumulative variance of the first six factors was 80.996%, greater than 80%, so it was appropriate to select six factors. Each index corresponds to 3 research variables, a total of 18, as shown in Table 2. In order to better obtain the research results, the influence degree of each factor is quantified through 1 to 5, 1 means the influence degree is small, and 5 means the influence degree is large.

**Table 1** Description of measurement factors

<table>
<thead>
<tr>
<th>Factor</th>
<th>Description of measurement factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group psychology</td>
<td>Group psychology1. My view of the behavior of others buying mystery boxes</td>
</tr>
<tr>
<td></td>
<td>Group psychology2. Someone around me buying mystery boxes will cause me to buy mystery boxes</td>
</tr>
<tr>
<td></td>
<td>Group psychology3. Social trends may drive me to buy mystery boxes</td>
</tr>
<tr>
<td>Symbolic consumption</td>
<td>Symbolic consumption1. I'd be inclined to buy certain ips from mystery boxes</td>
</tr>
<tr>
<td></td>
<td>Symbolic consumption2 &quot;I&quot; will be attracted to mystery boxes that offer rare or special items and buy them</td>
</tr>
<tr>
<td></td>
<td>Symbolic consumption3. The cultural and emotional sense of belonging brought by mystery boxes will make &quot;I&quot; strengthen the desire to buy mystery boxes</td>
</tr>
<tr>
<td>Independent consumption</td>
<td>Independent consumption1. After becoming a college student, my consumption is completely controlled by myself</td>
</tr>
<tr>
<td></td>
<td>Independent consumption2. The proportion of mystery boxes that I buy increases with the amount of money available</td>
</tr>
<tr>
<td></td>
<td>Independent consumption3. If I had my own independent income, I'd be willing to use it to buy mystery boxes</td>
</tr>
<tr>
<td>Fashion Consumption</td>
<td>Fashion Consumption1. &quot;I&quot; am a trendsetter</td>
</tr>
<tr>
<td></td>
<td>Fashion Consumption2. I think buying mystery boxes is trendy</td>
</tr>
</tbody>
</table>
4 RESEARCH PROCESS AND RESULTS ANALYSIS

4.1 Descriptive Statistical Analysis

The survey results show that nearly 90% of the respondents say that someone around them has purchased mystery boxes, with female consumers showing higher interest in mystery boxes. Most consumers purchase mystery boxes infrequently. More than 80% of mystery boxes consumers purchase mystery boxes 1-3 times per month, but there are still a few people who are addicted to Mystery boxes. College students don't spend much on mystery boxes, with only 3.5% of respondents reporting spending more than 1,000 yuan.

The top reason consumers buy is "mystery boxes opening surprise", with a weighting of 76.32%. In line with the public's curiosity psychology is the biggest marketing selling point of mystery boxes, its weighted proportion is up to 83.76%, in all the survey advantage attributes in a prominent leading position. Most of the respondents are satisfied with the mystery boxes, while many of them indicate that the price performance of mystery boxes is average or even low. Surprisingly, most respondents report that the actual product they get when they buy mystery boxes is not that different from what they expect.

In addition, the survey results show that 72.5% of people are willing to consume their favorite IP products, which indicates that most of the mystery boxes consumers are driven by symbolic factors. At the same time, nearly half said they were more fashionable. Thus, the foundation of the connection between Fashion Consumption and mystery boxes consumption is established.

4.2 PLS PM Model

This paper studies the complete multivariate correlation of the above six factors to construct a formative model that can effectively explain Perceived value of mystery boxes of college students' consumption and verify the validity of the research hypothesis. As can be seen from Table 1, the five influencing factors all have a significant impact on Perceived value (T>1.96). Among them, Fashion Consumption has the largest impact on post-purchase psychology (0.285), while Group psychology has the smallest impact on post-purchase psychology (0.136). The coefficients of all factors are greater than 0, that is, the hypothesis of this paper is valid.
### Table 2 Effect table of PLS-PM model

<table>
<thead>
<tr>
<th>Relationships</th>
<th>Path coefficient</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group psychology →</td>
<td>0.136</td>
<td>3.397</td>
</tr>
<tr>
<td>Independent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>consumption →</td>
<td>0.203</td>
<td>5.227</td>
</tr>
<tr>
<td>Fashion Consumption</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Symbolic consumption</td>
<td>0.285</td>
<td>6.424</td>
</tr>
<tr>
<td>Operate conditioning</td>
<td>0.183</td>
<td>4.034</td>
</tr>
</tbody>
</table>

#### 4.3 Apriori Association Analysis

Based on the 15 factors corresponding to the above five factors, the association rule mining method (Apriori) is used to mine the influencing factors that lead to a higher Perceived value in mystery boxes consumption. The top five association rules with the highest confidence can be obtained when 100% Perceived value has 100% probability to purchase mystery boxes and rule support is greater than 20.

People with characteristics such as "I will be attracted and bought by rare or special models introduced by mystery boxes products", "good empathy experience, that is, the cultural and emotional belonging brought by mystery boxes products will increase my desire to buy mystery boxes", and "If I have my own independent income, I am willing to use the money to buy mystery boxes" have the highest support among the top five groups.

#### 4.4 Logistic Regression

After the previous correlation analysis, the specific impact degree of each factor affecting the consumption intention of mystery boxes is further explored, and logistic regression is conducted. Next, Group psychology1, Group psychology2, Group psychology3, Symbolic consumption2, Symbolic consumption3, Independent consumption1, Independent consumption3 and Fashion Consumption1 were used to construct the model. The inspection information is as follows in Table 3 and Table 4

### Table 3 Summary of comprehensive test of model coefficients

<table>
<thead>
<tr>
<th>log likelihood</th>
<th>Cox &amp; Snell R</th>
<th>Nagelkerke R</th>
</tr>
</thead>
<tbody>
<tr>
<td>323.859</td>
<td>0.268</td>
<td>0.358</td>
</tr>
</tbody>
</table>

### Table 4 Omnibus test of model coefficients

<table>
<thead>
<tr>
<th>Chi-square</th>
<th>Degree of freedom</th>
<th>Significant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Steps</td>
<td>94.440</td>
<td>8</td>
</tr>
<tr>
<td>Block</td>
<td>94.440</td>
<td>8</td>
</tr>
<tr>
<td>Model</td>
<td>94.440</td>
<td>8</td>
</tr>
</tbody>
</table>
From the summary of the model, the Cox & Snell R value is 0.268, and the Nagelkerke R value is 0.358. In the study of practical problems, the R2 value is large, indicating that the model is effective in fitting the original data. Due to the limitation of the questionnaire, there may be more variables that can explain the dependent variable that are not considered. Finally, it can be seen from the Omnibus test table of the model that when the significance level is 0.05 and the number of degrees of freedom df=8, the chi-square value of the model calculated by SPSS is 94.440, which is greater than the critical value of 15.51 under this degree of freedom. Therefore, the model passes the test when the significance level is 0.05.

The dependent variable of model variables is defined as Y, which means Perceived value=1 (higher). The independent variable is X, which means perceived value of others buying mystery boxes =1 (higher) and other factors supporting the consumption of mystery boxes.

The eight factors involved in the five association rules are included into the independent variables, and logistic regression is conducted between them and purchase behavior. In these five association rules, the influence of each factor on purchase behavior can be found when other variables are the same.

In particular, when other variables are equal, "I" will be attracted by rare or special models launched by mystery boxes products and buy a higher degree of perceived value (=1) than "I" will be attracted to rare or special models launched by mystery boxes products and buy less likely (=0) 3.82 times; Perceived values that are willing to buy mystery boxes to a higher degree (=1) with their own independent income sources are 2.64 times more willing to use their own independent sources of income to buy mystery boxes to a lower degree (=0), etc.

4.5 Random Forest

Purchasing behavior is regarded as a dichotomous classified dependent variable, and the five factors mentioned above are used. The three independent variables under each factor are determined by the structure of the questionnaire. According to the statistical description analysis, gender factors and the attention paid to the promotion of mystery boxes have a significant impact on the result of purchasing mystery boxes, so there are 16 independent variables in total. Purchasing behavior is regarded as a dichotomous classified dependent variable, and the five factors mentioned above are used. The three independent variables under each factor are determined by the structure of the questionnaire. According to the statistical description analysis, gender factors and the attention paid to the promotion of mystery boxes have a significant impact on the result of purchasing mystery boxes, so there are 16 independent variables in total.

The results show that the degree of pursuit of fashion, the degree of autonomy in personal funds and the willingness to buy mystery boxes with their own independent income sources have significant effects on the purchasing behavior of mystery boxes fans, and the logistic regression results are verified.
5 STUDY THE CONSUMPTION BEHAVIOR OF MYSTERY BOXES FROM THE PERSPECTIVE OF BEHAVIORAL ECONOMICS

Behavioral economics is a discipline that studies economic phenomena under incompletely rational behavior. The introduction of psychology into economics increases the ability of economics to explain various economic phenomena in real life, thus providing a psychological cornerstone for economic analysis, making economic analysis closer to the reality of life and making policy formulation more reasonable.

5.1 Mystery Boxes Consumption under the Experience Economy

Experience economy, the extension of service economy, is the fourth economic type after agricultural economy, industrial economy and service economy. It emphasizes the satisfaction of customers' feelings and attaches importance to the psychological experience of customers when consuming behaviors occur [6]. Maslow(1943) divides human needs into five levels, namely physiological needs, safety needs, social needs, respect needs and self-actualization needs. Experience economy is the inevitable trend of the sublimation of human needs. As shown in FIG. 2, in the era of experience economy, consumers' consumption behaviors are shown as follows: From the perspective of consumption structure, consumers' emotional demand increases and their demand for personalized products and services becomes higher and higher; [7]

Today, mystery boxes, which license IP for upstream art and design, will lead the "fashionable toy IP economy." Today, mystery boxes, which license IP for upstream art and design, will lead the "fashionable toy IP economy." The introduction of limited style and collection style can accelerate the creation of IP, meet consumers' desire for collection and improve consumers' identity in the cultural circle where they are interested. On the other hand, the stimulation of mystery boxes satisfies the instinct of human beings to pursue uncertainty. Everyone has their own consumption principles. For Gen Z college students, when their self-actualization needs are linked to consumption, they will naturally pursue individuation and differentiation in products and services to satisfy their higher emotional needs, reflecting the mystery boxes are products of the experience economy.

5.2 Irrational Consumption under Mental Accounting

Due to the existence of consumer Mental Accounting, consumers tend to violate some simple economic algorithms when making decisions, thus making many irrational consumption behaviors. Mental Accounting is a fictional account that classifies and counts different consumption items according to the mental cognitive system and makes irrational economic
decisions. It will change the psychology according to the special changes of transaction items, and it will be guided and changed by the mental cognitive system. For example, no matter how much you spend on food that goes over your budget, it won't affect your desire to buy mystery boxes that day. This is the effect of placing the budget for eating in a different mental account than the budget for buying mystery boxes.

![FIG. 3 Psychological accounts of college students](image)

The disposable income of college students is divided into what their parents give and what they earn. Most of the unpaid income comes from parents. As for this part of income, college students can not reasonably estimate their psychological accounts because they have not paid their labor. If the various consumption items into the psychological account at random, resulting in irrational consumption.

6 CONCLUSIONS AND SUGGESTIONS

Experience economy is the inevitable trend of human needs upgrading. It is true that college students purchase mystery boxes to some extent to pursue personalized experience and satisfy emotional needs brought by materials. Every purchase of mystery boxes by a consumer is a psychological game against the enterprise. The joy of uncertain returns brought by buying unknown goods makes consumers happy in the "self-realization", the re-shaping of the value creation subject and the psychological pulse changes through memory utility, and experience the pleasure of opening the blind box.

College students' purchase of mystery boxes is a hedonistic act used to satisfy the psychological needs of the self in the process of consumption. The consumer needs of consumers and the attributes of mystery boxes are connected, making mystery boxes consumers willing to provide them with a separate and irreplaceable psychological account, hoping to get self-satisfaction in the consumption process of mystery boxes.

What has the greatest influence on the mystery boxes Consumption behavior of college students is Fashion Consumption, while the least influence is Group psychology. The exciting gameplay of mystery boxes and its products derived from the creation of IP are the most
attractive to consumers. College students' overall satisfaction with Mystery Boxes products is not low. College students who pursue social trends are more likely to buy mystery boxes.

To a certain extent, mystery boxes purchased by college students are indeed for the pursuit of personalized experience and the satisfaction of emotional needs brought by materials, but there is the possibility of irrational consumption based on psychological accounts.

For college students, they should look at the popularity of mystery boxes rationally and do not follow the trend of consumption, so as not to be hyped by stakeholders. For businesses, developing IP culture is an effective way to increase user engagement. Only with a variety of IP can users continue to buy. However, merchants should regulate their own behavior, so as not to consume consumers' preference for mystery boxes. For the regulatory authorities, they should standardize the operation order of mystery boxes, make it clear that "mystery boxes are used for entertainment rather than for speculation", and crack down on "scalpers" and fake goods according to law. For the regulatory authorities, they should standardize the operation order of mystery boxes, make it clear that "mystery boxes are used for entertainment rather than for speculation", and crack down on "scalpers" and fake goods according to law.

REFERENCES


Research on the Realization Model of Accounting Information Sharing

Yicai Zhu¹,a*, Jilin Qu¹,b

*zxpmwzyc@163.com¹, qujilin@126.com¹

Shandong University of Finance and Economics Jinan, Shandong, China¹

Abstract: Accounting information sharing plays an important role in fully leveraging the resource allocation effect of accounting information and improving economic management. In order to achieve accounting information "counting out one door and sharing resources", this article studies the implementation mode of accounting information sharing. Establish the platform of accounting information, form a vertical connectivity, horizontal interaction, internal integration, national unity, and open and transparent sharing service system of accounting information, and achieve "one platform, five level application" of accounting information. Provide specialized accounting information to government departments through a business integration model. Provide public accounting information to the public through information retrieval mode. Cloud computing is a way of providing users with large-scale and intensive services by treating resources, services, and applications as public facilities. Utilize cloud computing technology to achieve the integration of accounting information resources.

Keywords: Accounting Information Sharing, Accounting Information Sharing Model, Accounting Information Platform, Cloud Computing Technology.

1 INTRODUCTION

Accounting information is important foundational information that reflects economic activities. Accounting information is the guide of resource allocation, a guide of the capital market, and the main carrier of economic information. It plays a significant role in economic management [1]. However, for a long time, there have been many problems in accounting information disclosure and sharing in China. On the one hand, government departments are unable to share accounting information, each acting independently, with different standards, and multiple submissions, making it difficult for regulatory authorities to effectively integrate. On the other hand, the relevance, timeliness, and completeness of publicly disclosed accounting information are poor, which cannot meet the requirements of information users.

In order to change the situation where accounting information cannot be shared, the issue of accounting information sharing has been explored since the 1980s. The Ministry of Finance has proposed the goal of "counting out one door and sharing resources" for accounting information, and has always regarded accounting information sharing as the main goal and overall task of accounting reform and development. In 2016, the Ministry of Finance issued the Outline of the 13th Five Year Plan for Accounting Reform and Development, proposing the establishment of a socialized accounting information public service platform primarily
focused on disclosing financial report data disclosure [2]. But after nearly forty years of exploration, accounting information sharing is still in its infancy.

The significance of achieving accounting information sharing is significant. Firstly, achieve interconnectivity and information sharing among government departments. Promote a good combination of various regulations and fully suppress accounting information distortion. Secondly, effectively leverage the resource allocation effect of accounting information, improve the macroeconomic regulation system, and make macroeconomic regulation more scientific and effective. Thirdly, as an important component of financial work, promote the integration and construction of financial business systems, promote horizontal integration, vertical centralization, and national systematization, and improve financial governance. Fourthly, provide simple and easy-to-use accounting information to the public and other accounting information users to meet their accounting information needs. Strengthen social management and supervision, and effectively leverage the management role of accounting information. This article studies the implementation mode of accounting information sharing.

2 ACCOUNTING INFORMATION SHARING MECHANISM

2.1 Sharing Method of Accounting Information

According to the goal of "counting out one door and sharing resources" in accounting information, accounting information platforms will be established at all levels of financial departments at the county level to comprehensively collect accounting information from various units within the region. By reviewing, summarizing, and analyzing, and collecting, storing, and reporting information level by level, establishes county-level, city-level, provincial-level, and national accounting information platforms. Establish comprehensive accounting information sharing service system with vertical connectivity, horizontal interaction, internal integration, national unity, and transparency. Implement "one platform, five level application" for accounting information [3]. See Figure 1.

A platform refers to a national accounting information platform. Taking county-level accounting information platforms as nodes gradually integrate accounting information resources and establish a unified national accounting information comprehensive service platform. The fifth level application refers to the use of accounting information platforms to achieve the sharing and application of accounting information among government departments at county, prefecture, city, provincial, and national levels, as well as the public. Through vertical integration, achieve vertical accounting information sharing among government departments at all levels of the country, province, prefecture, and county. Through horizontal interaction, achieve horizontal accounting information sharing among various government departments in the region. Utilize the Internet to provide information services to the public.
Firstly, through vertical integration, the accounting information of the financial department can be collected and shared step by step. Taking counties as nodes, comprehensively collect accounting information from various units within the region. Review, summarize, and analyze accounting information. Submit the information to the superior financial department to achieve level by level collection and vertical sharing of accounting information. By vertically connecting, establish accounting information databases at all levels of financial departments to address issues such as decentralized management and non-sharing of accounting information. Ensure that accounting information is relevant, reliable, and complete through data review.

Secondly, through horizontal interaction, achieve the interconnection and exchange of accounting information between the financial department and other relevant government departments. The accounting information required by relevant government departments is uniformly provided by the finance department and extracted directly from the accounting information platform. Achieve the goal of "counting out one door and sharing resources" of accounting information within the region, and solve the problems of multiple submission of accounting information, different caliber, decentralized management, and independent governance. Achieve cross departmental regulatory information sharing and promote effective integration of various regulatory measures.

Thirdly, through internal integration, achieve the integration of accounting information and related business within the financial department. The accounting information required for various internal businesses of the financial department is provided uniformly by the accounting information platform, such as budget management, financial supervision, centralized treasury payment, and other information, to achieve the integration of accounting
information and related businesses, and promote the integration and construction of business systems. The problem of internal business separation within the finance department is solved. The internal accounting information of the financial department to be "counting out one door and sharing resources" is realized, to enhance the scientific and effective nature of macroeconomic regulation decisions, and improve the level of financial governance.

Fourthly, through national unification, meet the needs of accounting information users such as the public. Establish a national unified standard system for sharing accounting information, including compilation requirements, data structure, and storage format. Solve issues such as data standards and formats to facilitate the acquisition, exchange, and sharing of accounting information. On the basis of sharing accounting information among government departments, other accounting information users such as the public can search for accounting information through the internet. Meet the needs of accounting information users such as the public, improve the efficiency of social management and public services, and improve the effectiveness of accounting information governance.

Fifth, through openness and transparency, make accounting information easily accessible to the public and other information users. Obtain accounting information through various means such as information retrieval through websites and other channels, and obtain diversified and personalized services. Relevant government departments release accounting information quality inspection announcements and other information to the public through accounting information platforms.

2.2 Service Mode of Accounting Information Sharing

2.2.1 Users of Accounting Information and Their Demands

According to the characteristics of the users of accounting information, the users of accounting information are divided into three categories: financial departments, government departments, and other stakeholders such as the public. Accounting information sharing needs to serve economic management and serve them simultaneously. The finance department is responsible for managing finance and accounting work and is the main user of accounting information. The relevant government departments include tax departments, market supervision and management departments, securities regulatory departments, state-owned asset supervision committees, and statistical departments. They use accounting information according to their respective functions as the basis for macro management and regulation. The public includes investors, creditors, business managers, employees, suppliers, customers, intermediaries, and industry associations. They have the right to understand the business operation status of the business. Government departments such as finance are relatively concentrated, while the public is more dispersed.

According to the requirements of different accounting information users, we divide accounting information into public information, special information and interactive information.

a) Public information. This includes basic information such as accounting statements, as well as public information released by various government departments, for the public and all accounting information users to access.

b) Special information. According to their own management functions, government departments use relevant information for market regulation and macroeconomic regulation,
such as tax payment information from tax authorities, for internal use by relevant government departments. Specialized information includes both original accounting information and analytical information. Analytical information refers to the comprehensive information obtained by summarizing and analyzing enterprise accounting information, mainly providing decision-making support for macro management and financial departments.

c)  *Exchange information*. The relevant information disclosed by government departments for other government departments to share such as the quality inspection of accounting information disclosed by financial supervision departments, is directly accessible to securities, auditing and other relevant departments, promoting a good integration of supervision and increasing the effectiveness of macroeconomic regulation.

### 2.2.2 Service Mode of Accounting Information Sharing

According to the characteristics of the needs and usage permissions of accounting information users, different sharing modes are adopted to meet the needs of accounting information users. Provide specialized accounting information to government departments through a business integration service model. Provide public accounting information to the public through a search model [4]. See Figure 2.

![Figure 2. Service mode of accounting information](image)

a)  *Business integration service pattern*. The relevant business systems of government departments such as finance are directly integrated with the accounting information platform to achieve information interconnection and exchange. Share accounting information through the XBRL standard data interface to ensure accurate, timely, and complete information sharing.

b)  *Information retrieval service pattern*. The public, including intermediaries, use information retrieval methods on websites to obtain accounting information.

To facilitate the use of different information users, accounting information sharing is divided into vertical sharing, horizontal sharing, internal sharing, and social sharing. See Figure 2.

Vertical sharing enables government agencies to share accounting information between superiors and subordinates. Vertical sharing is divided into two methods: top-down and bottom-up. The higher-level government departments use the accounting information submitted by the lower level government departments in the region to conduct macro regulation and decision-making. Lower level government departments utilize the comprehensive information released by higher level government departments to understand and grasp the comprehensive situation of the local area within the same level region.
Horizontal sharing enables accounting information to be shared among government departments at the same level. For example, the audit department and the finance department share accounting information with each other. Different government departments have specific work; have their own needs for accounting information. Currently, there are problems with multiple submissions and varying standards of accounting information in China, and there is a lack of information sharing among government departments. The main problem that needs to be solved in information sharing is to achieve horizontal sharing and interconnection of government departments' information, and to enhance the synergy of macroeconomic regulation.

Internal sharing enables the integration of accounting information with relevant business within government departments, and the sharing of relevant business within departments, also known as business integration. Accounting information is an important basis for macro management and regulation by the government departments concerned, and is the fundamental information for internal business operations of the government departments concerned. Within the financial department, budget management, centralized treasury payments, and financial supervision are all directly related to accounting information.

Social sharing enables the public and other information users to easily access accounting information, achieving transparency in accounting information disclosure. On the one hand, improve fiscal transparency and effectively leverage the governance effect of accounting information. On the other hand, build a service-oriented government to serve the construction of a market economy.

3 RESOURCE INTEGRATION OF ACCOUNTING INFORMATION SHARING

Accounting information can only be shared by collecting and integrating it. Compared to listed companies, accounting information platforms have a much larger scope of information disclosure. The integration method of utilizing accounting information resources of listed companies cannot meet the functional requirements of sharing accounting information. This article proposes a method for integrating accounting information resources based on cloud computing.

3.1 Overview of Cloud Computing Technology

3.1.1 Concept and Characteristics of Cloud Computing Technology

In the traditional information technology model, users not only need to purchase hardware, software and other infrastructure, but also need to be equipped with specialized software and hardware maintenance personnel. If a service provider can provide software and hardware rental services, it not only saves the cost of purchasing software and hardware facilities, but also avoids complex hardware and software maintenance.

Cloud computing is a way of providing users with large-scale and intensive services by treating resources, services, and applications as public facilities. As long as users are connected to the internet, they can use hardware, software, data and other resources on the "cloud" end at any time. Narrowly speaking, cloud computing is a network that provides
resource services, allowing users to access resources on the "cloud" end at any time. Just like a power plant, users can use electricity at any time without having to purchase their own generators and other facilities. In a broad sense, cloud computing is an internet-based hardware, software, and information service, and the shared pool of such service resources is called "cloud". Cloud computing integrates many resources for unified management, scheduling, and control, providing services to users [5].

The core idea of cloud computing technology is to integrate hardware, software, and data resources dispersed in different geographical locations through network connections, forming a virtualized computer resource pool for unified management and scheduling, providing services to users, enabling them to use them on demand like public service facilities such as water and electricity [6].

At present, cloud computing has been applied in many fields, such as Baidu Cloud, email, Alibaba Financial Cloud, and so on. The services provided by cloud computing to users are mainly divided into three types, including Infrastructure as a Service IaaS, Platform as a Service PaaS, and Software as a Service SaaS. The main advantages of cloud computing are cost saving, strong scalability, resource sharing, Storage security, and ease of use.

### 3.1.2 Security of Cloud Computing

To meet users' requirements for information security, cloud computing provides three different deployment solutions, including public, private, and mixed clouds.

a) **Public cloud.** Public cloud is built by service providers to build infrastructure, integrate cloud virtual resource pools, and provide services to multiple users. Users do not need to purchase complex devices and can use them through the internet. The main characteristic of public cloud is its low user cost, which is suitable for users who unconditionally set up private clouds or have low requirements for software and hardware environments. Users do not need to purchase various complex hardware and software, nor do they need to consider issues such as the security of the operating environment. Through the Internet, they can dynamically, flexibly, and self-service access resources in the public cloud, with high cost-effectiveness. However, public clouds are managed and controlled by service providers, and users cannot directly control cloud resources. Public clouds usually cannot meet many security regulatory compliance requirements, and may not be suitable for users with high requirements for information security and confidentiality.

b) **Private cloud.** Private clouds are built by service providers or units themselves to provide services specifically for a particular unit. Private clouds can be built by the unit itself, deployed within the unit or in a hosting location, or built by service providers. The resources in the private cloud are only shared within the unit and cannot be accessed by external users. Private clouds can be managed and operated by units themselves or service providers, and users have high control over cloud resources. The security, confidentiality, and service quality of data can be better guaranteed, making them suitable for applications with high data security requirements. However, all software, hardware, and other facilities in private clouds need to be purchased by one, and maintenance personnel need to be equipped, resulting in relatively high construction costs.
c) Mixed cloud. Mixed cloud is a combination of public and private cloud services. In cloud computing technology, users can build their own private cloud and also use the public cloud provided by service providers. The combination of these two service methods constitutes mixed cloud. A mixed cloud combines the advantages of both public and private clouds.

3.2 Basic Ideas of Accounting Information Resource Integration under Cloud Computing Technology

Cloud computing technology integrates software, hardware, and data resources dispersed in different geographical locations through network connections for unified management and scheduling, providing services to users. This service method provides an effective way for the integration of accounting information resources. The basic idea for integrating accounting information resources under cloud computing technology is as follows.

3.2.1 Distributed Storage, Centralized Management

The resources of accounting information sharing come from grass-roots enterprises and are distributed in all counties and regions of the country. Accounting information is widely distributed and has a large amount of information. It is very difficult to collect and integrate it in a certain time.

Cloud computing technology integrates hardware, software, and data resources dispersed in different geographical locations through network connections for unified management and scheduling, providing services to users. This service method provides an effective way to integrate accounting information resources. The basic idea for integrating accounting information resources under cloud computing technology is as follows.

With cloud storage technology, enterprise accounting information can be stored in provincial financial departments without all transmission and storage being centralized. By establishing an accounting information foundation cloud in various provinces, it will be integrated into a national accounting information cloud. On the one hand, it avoids various problems caused by centralized data transmission and storage, and improves information processing efficiency. On the other hand, it is conducive to the horizontal sharing of accounting information within the province. For some developed regions, accounting information can also be stored in prefectures to better improve information processing and sharing efficiency.

3.2.2 Unified Scheduling and Resource Sharing

Cloud computing integrates data resources distributed in different geographical locations and devices, manages and schedules them uniformly, adopts unified information security and sharing service strategies, and provides accounting information sharing services for different users. Users don't need to worry about where the data is stored; they can share accounting information from enterprises across the country through the internet. This approach avoids issues such as data inconsistency and information security caused by decentralized data management, and provides guarantees for achieving the goal of "counting out one door and sharing resources".
3.2.3 Centralized Control to Ensure Safety

Cloud computing provides three different deployment solutions: public cloud, private cloud, and hybrid cloud, providing technical support to ensure the security and confidentiality of information.

According to the goal of accounting information sharing, accounting information resource sharing in the cloud computing environment adopts a distributed storage, unified management, centralized control, and decentralized service approach. The accounting information platform adopts a hybrid cloud solution, where private clouds mainly provide services to relevant government departments, while public clouds mainly provide services to other stakeholders such as the public to ensure control of information access permissions and improve information security and confidentiality.

According to the management functions of government departments, private clouds can be further divided into financial department private clouds and government related department private clouds. Among them, the private cloud of the finance department mainly provides services for the finance department, including basic accounting information of enterprises and public institutions, as well as relevant information required by various business systems within the finance department. The private cloud of government departments mainly provides services for government departments, including relevant business information required for their management functions.

3.2.4 Dynamic Expansion to Reduce Investment

Cloud computing can achieve dynamic and scalable scalability, not only adding hardware facilities according to user needs, but also expanding functions and services.

The construction of an accounting information platform is a complex project, and the implementation of information sharing functions needs to be simplified and gradually improved. Therefore, the scalability of cloud computing provides favorable conditions for the overall design and step-by-step implementation of accounting information platform construction.

Cloud computing technology can fully utilize the existing hardware and software infrastructure of the financial department. On the one hand, it can save costs and reduce investment, and on the other hand, it is conducive to improving construction efficiency.

3.3 Implementation Method of Accounting Information Resource Integration under Cloud Computing Technology

For the purpose of "counting out one door and sharing resources" in accounting information, according to the sharing mechanism of accounting information, the construction of accounting information platforms takes the county as the node and the province as base. Cloud computing technology is used to gradually integrate resources of accounting information, and four levels of accounting information platforms are established at the county, city, province, and national levels to provide services for accounting information users [7]. As shown in Figure 3.
The accounting information platform in the cloud computing environment is mainly composed of four parts: user service layer, resource management layer, resource integration layer, and security layer.

### 3.3.1 Support Layer

The security layer mainly provides security, legal, standard, and policy guarantees for accounting information sharing, including policy guarantees, coordination mechanisms, and security guarantees.

**a) Policy guarantee.** The information shared by accounting information comes from enterprises, with information users including relevant government departments and the social public. The collection and sharing of information cover a wide range. Therefore, it is necessary to formulate sound policies in order to achieve the purpose of accounting information sharing. The main content includes enterprise accounting information reporting mechanism, review mechanism, service mechanism, security guarantee mechanism, and supervision mechanism.

**b) Coordination mechanism.** Government departments have different requirements and responsibilities for accounting information. Accounting information sharing and sharing have changed the traditional channels and utilization of accounting information. Therefore, it is necessary to establish a coordination mechanism between government departments, including the sharing content and transmission methods of information between departments, in order to achieve the interconnection and exchange of accounting information between government departments and achieve the goal of accounting information sharing.
c) **Security guarantee.** To ensure the security of accounting information, information security measures are formulated from both institutional and technical aspects, including user permission control, data storage management, information security monitoring system, and emergency response measures.

3.3.2 **Resource Integration Layer**

The resource integration layer is responsible for the integration of resources of accounting information. Mainly collecting and integrating enterprise financial report information [8]. The accounting information of enterprises is submitted to the county-level finance department, which integrates the accounting information of counties, cities, and provinces step by step, and migrates to a virtualized environment to form a "provincial-level accounting information cloud". The fusion of various provincial-level accounting information clouds forms a "national accounting information cloud".

The resource integration layer is mainly divided into two parts: data integration and data virtualization. Data integration is responsible for reviewing the accounting information submitted by enterprises, converting the format according to the XBRL standard specifications, and then transmitting, integrating, and storing it level by level. Data virtualization aggregates various distributed resources under cloud computing technology and stores them in cloud computing resource pools, preparing for the use of information resources.

3.3.3 **Resource Management Layer**

The resource management layer mainly includes resource management and task allocation for cloud computing. Resource management is responsible for balancing cloud resource nodes, preparing user services, and determining which information resources enter the public cloud service system and which information resources enter the private cloud service system to ensure the security of accounting information resources.

Task deployment is responsible for scheduling and executing resource use tasks submitted by users, so that resources can efficiently provide services to users.

3.3.4 **User Service Layer**

The user service layer mainly provides services to users and is responsible for user management and access control. Through identity authentication, determine the type and usage permissions of users, provide corresponding services according to different service modes, and control the content accessed by users through access permissions to ensure the security of information.

a) **Financial department.** Accounting information is the basic information of the relevant business of the financial department, including budget management, financial supervision, and so on. The platform of accounting information provides services for the financial department through business integration. The accounting information required for the relevant business of the financial department can be obtained directly from the platform of accounting information through data interfaces, achieving the integration of accounting information and related businesses.
b) Relevant government departments. Relevant government departments. The platform of accounting information provides services to government departments through a business integration service model. The accounting information required for government departments’ related businesses is directly obtained from the accounting information platform through data interfaces, achieving the interconnection and sharing of accounting information between financial departments and government related departments.

Government departments such as finance can publish relevant business information to private clouds for sharing among other relevant government departments, achieving regulatory information sharing among government departments.

Relevant government departments such as finance can also release public information related to regulation to the public cloud for the public and other information users to access. For example, the Ministry of Finance regularly issues accounting information quality inspection announcements, etc.

c) The public. The platform of accounting information provides services to the public and other accounting information users through the service model of information retrieval. The public can access public information in the public cloud through the internet through devices such as computers and mobile phones.

4 IMPLEMENTATION AND APPLICATION OF ACCOUNTING INFORMATION PLATFORM

The establishment and operation of the platform of accounting information should be led by the Ministry of Finance, with the participation of financial departments at all levels and the cooperation of other government departments. The market can play a supporting role. The construction of the platform of accounting information should be planned as a whole, implemented step by step, based on provinces, piloted first and then promoted, simplified first and then complex, and gradually improved. Develop relevant policies and regulations to ensure the construction and operation of the platform of accounting information. Coordinate the relationship between accounting information platforms and enterprises. Develop corresponding software.

We should deeply tap the value of accounting information resources, improve the utilization of accounting information resources, and better provide accounting information for government departments and the public.

5 CONCLUSION

Accounting information sharing is of great significance for fully leveraging the resource allocation effect of accounting information and improving economic management. This article studies the implementation mode of accounting information sharing to achieve accounting information "counting out one door and sharing resources". Establish an accounting information platform to achieve a "one platform, five level application" of accounting information, forming a vertical connectivity, horizontal interaction, internal integration,
national unity, and open and transparent accounting information sharing service system. Provide specialized accounting information to government departments through a business integration model. Provide public accounting information to the public through information retrieval mode. Utilize cloud computing technology to achieve the integration of accounting information resources.

Acknowledgement: This study is supported by the National Social Science Fund Project “Research on Accounting Information Sharing Mechanism from the Perspective of National Governance”, No: 15BGL057.

REFERENCES

http://kjs.mof.gov.cn/zhengcefabu/201610/t20161018_2437976.htm
https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&dbname=CJFDLAST2019&filename=CZ YJ201811013&uniplatform=NZKPT&v=MsFejWH_VOpBoRyHydbkQOen3cnQyQ68L_cqofEB D9v5V0_0TjOR_tk10aq9
The Role of Information and Communication Technologies in Modern Business Management

Zhian Ye\textsuperscript{1}, Natalia Karmina\textsuperscript{2}

1032208771@rudn.ru\textsuperscript{1}, nkarmina@mail.ru\textsuperscript{2}

Department of National Economy, Peoples' Friendship University of Russia (RUDN) Moscow, Russia, Moscow, Russia\textsuperscript{1}
Senior Lecturer of the Department of Advertising and Business Communications, Peoples' Friendship University of Russia (RUDN) Moscow, Russia\textsuperscript{2}

Abstract: The last few decades have been characterized by a rapid saturation of data processing systems in almost all spheres of society. Information has become a strategic resource of society, which has led to the inevitable spread of information and communication technologies. The article deals with the concepts of information and communication technologies in business management. Their role and the need for application in modern business have been investigated. Conclusions are made about the need to solve the problems of the formation of information and communication technologies and their role in management systems. In this article, information technology means methods of solutions for the implementation of running and supporting a business, by creating a system for collecting, storing, processing and transmitting the necessary information.

Keywords: Information Technology, Information, Communication, Technologies, Business Management.

1 INTRODUCTION

Today, in a competitive market environment, every company must have the ability to adapt quickly and economically to constantly changing market conditions. The emergence and development of information technology (IT) has had a significant impact on the development of all sectors of the economy. The information society began to take shape at the turn of the late 20th and early 21st centuries. The growth of the IT industry occurred in the late 1990s due to the development of the Internet and investment in infrastructure. Information has become one of the main factors in the development of production. At the beginning of their development, information and communication technologies (ICTs) had a high cost and were more often used for experimental purposes. Every year, information technologies are becoming more accessible and of higher quality, and their scope of application is expanding. The social needs of consumers have increased, but there are also new ways to meet these needs.

The proof is that the use of ICT is already giving growth to the Russian economy: there is unlimited access to goods and services; the labor market is changing; the competitiveness of Russian goods on world markets is increasing; the efficiency of inter-industry and interregional cooperation within the country is increasing; virtual money, banks and payment systems are
appearing. The Internet allows you to save time and money of Russian citizens on purchases, since goods in online stores are cheaper, and you can manage your money accounts without leaving your home[3].

This article examines the economic sphere and the impact of information technologies on it. In this context, information technologies are methods of solving problems of management and business support by creating a system for collecting, storing, processing and transmitting the necessary information. The essence of the methods is a single result obtained with a complex approach of interacting elements.

The main meaning of IT work is the implementation of program actions aimed at obtaining a specific result using information systems (IS). Most often, this result is the provision of information to a specific consumer and preparation for the next action. The process of collecting, processing, storing and transmitting information must exclude possible errors.

The use of ICTs is promising, and therefore the political state task in the field of economics is to expand their use:
- improving information security due to independence from imports;
- introduction of ICT in the construction sector [3];
- development of e-commerce;
- development of the system of public administration, economy, business, social sphere and the whole society through the program "Digital Economy" [5].

2 METHODS, RESEARCH PROCESS

In the study, the authors used some methods, such as analysis and synthesis, induction and deduction, historical and logical, abstraction and refinement.

The research process is aimed at determining the role and significance of information technology in the modern management of a company.

3 RESULTS

In some areas, technology is critical for business, enabling effective management of the point-of-sale system, implementing the use of ICT in general business management, building accounting systems, and other complex aspects of the firm's day-to-day operations. Even such a simple thing as a calculator, revolutionary in its time, appeared thanks to technology. It's hard to imagine how you can go back to performing tasks manually. That would take us back a hundred years.

The technology allows us to automate many processes, which increases our productivity, allows us to use less resources to improve the quality and speed of providing services to customers. Additionally, the introduction of new technologies makes it possible to serve more customers at a lower cost.
The technology also makes it easy to store more information while maintaining the integrity of that information. The quality of storing confidential information is improved, so that it becomes less vulnerable to data hacking. In this case, the information can be obtained instantly, when it is needed, and it can be analyzed not only to study past trends, but also to predict the future. As a result, ICTs help businesses, improve the quality of strategic and operational-tactical management decisions.

Communication is part of the business. Thus, transport and processes turn a business into a network of complex processes that interact with each other. Thanks to technology, it has become possible to globalize business operations. Now almost everyone can do business almost anywhere in the world, from any room of their home.

Technology has enabled businesses to have a broader reach and reach global markets. The best example of this is the Internet. The internet is now an important part of any business’ marketing campaign, as it allows the business to attract customers around the world.

Technology, being well integrated into the business, has allowed us to make life more worthwhile. However, it would be foolish to deny that there are also threats to business caused by technology. These include malicious actions by various organizations, such as hacking. Because of this, it is important that businesses take responsibility when using technology to do business. With the advantages that technology brings, there are also disadvantages that also need to be addressed. Still, it's worth it, and we need to recognize and use technology responsibly to make our business better. Mobility is seen by many as the next big frontier for business. Google's algorithms reflect this, as they make mobile sites a priority. Your business and all aspects of it can be handled with just a tablet or smartphone-from content marketing to customer relations, from sales to internal organization issues like billing and delivery.

But mobile solutions are not just about business, they are also useful for consumers. The modern generation uses their phones for everything from buying and selling to sharing experiences with friends and finding local businesses.

Cloud computing has allowed companies to transfer their functions to third parties over the Internet. This makes it possible to handle packets with variable data, and also allows companies to quickly expand and leverage mobility without worrying about things like crashes, downtime, and data loss. All this has allowed small and medium-sized businesses to access resources that would have cost them a fortune a few years ago.

As more and more data becomes available, it is now much easier to analyze and gain a deep understanding of what customers are looking for. Analytics services are expanding every day and allow companies to divide their potential customers into more and more specific groups, which makes it much easier to target them and get a greater return on their advertising money costs. Something as simple as having a Google account can tell the company where the user is from, what browser they use, how they got to the site, what they do on that site, how long they can stay, and at what point they decide to leave. There are even more advanced analytics services that allow companies to further refine this segmentation to significantly improve their conversions. However, the impact of digitalization varies across countries and sectors. Developed countries benefit more from economic growth of almost 25%, although they tend to lag behind developing countries in job creation by the same margin [5]. The main reason for the differences in the effects of digitalization is the economic structures of developed and
developing countries. Developed countries rely mainly on domestic consumption, which attaches great importance to non-tradable sectors.

In all developed countries, digitalization increases labor productivity and has a tangible impact on economic growth. However, the result could be job losses, as lower-skilled jobs with lower added value are sent abroad to emerging markets where labor is cheaper. In contrast, emerging markets are more export-oriented and move at the expense of tradable sectors. They tend to benefit more from the impact of digital technologies on employment than from their impact on economic growth [4].

First, they must develop digitalization plans for the target sectors in which they want to maximize the impact of digitalization. Second, they should encourage the development of the necessary capabilities and tools to implement these digitalization plans. Finally, policy makers should work with industry, consumers and government agencies to create an inclusive information and communication technology (ICT) ecosystem that will facilitate the wider dissemination and use of digital services.

Today, digitalization is becoming a new tool for creating and maintaining absolute business advantages, and in some cases directly provides a victory over competitors in certain sectors, which is the most important opportunity underlying all other national economic efforts. Digital technologies have become an integral part of organizations’ activities. Currently, the vast majority of them work using one or another type of network technologies, software systems. At the same time, the degree of integration of digital tools into business processes varies in different countries due to differences in the level of infrastructure development and the sectoral structure of the economy. To measure the overall level of demand for key ICTs in the business sector, specialists from the National Research University Higher School of Economics (HSE) calculate the Business Digitalization Index (Fig.1).

Today, the digitalization of all processes in the company gives a great advantage, which is that everything becomes faster, better and cheaper. Digitalization of the business contributes to the efficiency of its processes, consistency and quality, namely (Fig.2):
Digital solutions provide a competitive advantage in the business world by improving quality, inventing new profitable methods, and ensuring consistency. Consequently, there are many companies that have adopted and are significantly increasing the digitalization of their business.

The process of informatization of the world economy follows the path of development of its two components: material (computer technologies, peripheral and telecommunications equipment) and non-material (the impact of information and informatization on the economy, software and information support). The information economy is based on scientific and technical information, as well as financial, investment, economic, political, educational, forecast and indexed information (Fig. 3).

A variety of information is necessary for the preparation of short-and long-term business plans and strategies. Information flows in various firms, through their influence on international
financial markets and the movement of investment resources, influence the dynamics of global economic relations. The enterprise should be considered as an information node, where external incoming information flows converge and internal flows circulate, and ICTs can increase the processing speed of constantly growing information flows. At the same time, as some researchers correctly point out, "the main economic effect of using these tools is to improve the quality of management and the quality of the main production processes, and not in direct savings from speeding up data processing."

Information about the state of affairs in the real sector allows you to adjust the plans and decisions made, as well as the strategies used. The large gap between the forecasts and the realities of economic activity shows that there is a threat of a recession or a state of crisis. Due to the importance of information at the present stage, there is an increase in investment in information, the demand for which is constantly growing. Information has a huge impact on business at all levels and at all scales: from small to large, from national to global, contributing to the competitiveness of firms in various sectors, so information is an attractive commodity for firms seeking to conquer new markets.

The information industry has also contributed to the emergence of many new firms and companies engaged in the production of computers, peripherals and components, fax equipment, mobile phones, information services, databases and knowledge, software. The emergence of the information industry has led to significant changes in the patterns of international trade and competition. Many of the above factors have inevitably led companies to reorganize their businesses and build their information infrastructure.

ICTs enable companies to achieve greater accuracy, speed, and efficiency in concluding, confirming, and executing commercial transactions. Thus, e-commerce combines all forms of business transactions and transactions carried out in electronic form. E-commerce issues have taken a strong place in the system of the main proportions of the world's leading economic countries. In the course of the development and promotion of the process of globalization, the computer industry, the information services, databases and knowledge sectors have undergone radical changes. The most important result of the widespread introduction of ICTs in all areas of economic and financial activity has been a fundamental change in the management system of a transnational corporation (TNC), for which mass processing of information, its transmission and use have become economically profitable, and the costs of it are economically justified.

The information economy provides economic growth in developed countries at the present stage. The information type of economic growth is mainly indicative for economically developed countries, the use of the information factor leads to an increase in the organization and order of the business environment. The information saturation of the ICT sector of the economy contributes to the increase in labor productivity. At the same time, the cost-performance ratio is steadily declining. The information industry produces information products and services. The information product has an intangible nature, which creates a problem of understanding the nature of information ownership, determining its value, as well as protecting intellectual property rights. It is interesting to note that even in the case of advertising on the Internet and when there are several price models for calculating the cost of advertising, this is not easy to do. With regard to information ownership, a significant amount of information is freely available to users. And here there are many problems: the right of ownership of scientific and creative activities, that is, the right to work in the field of science, intellectual property
rights, the cost of trade secrets of production or special knowledge. The cost of inventions and industrial designs. The information economy in developed countries is a new long cycle of economic development based on the production and sale of information products and services.

Thus, the use of ICT and network technologies in business increases the efficiency of internal and external activities of companies, ensuring faster interaction between suppliers and customers, as well as optimizing the synthesis of supply and demand.

The business world changed forever when computers came along. Businesses can use information technology, computers, and various software tools to work more efficiently across all departments, including finance, manufacturing, human resources, and security. The digital transformation model is shown in Figure 4.

![Digital transformation model](image)

**Fig. 4. Digital transformation model**

The digital transformation model allows all participants in business processes to see the full picture of what is worth paying attention to, and shows the sequence of actions. However, they do not give clear instructions, but only set benchmarks by which leaders can carry out transformations in companies, taking into account the current circumstances and specifics. Models on digital transformation show people at different levels of the organization the general trend and direction where they need to direct their efforts in order for digital transformation to be successful.

Thus, the digital transformation model allows:

- see the big picture ideal for leaders;
- create a list of tasks and determine the sequence of their implementation;
- focus on divisions responsible for sales, customer interaction and operational functions;
- adapt the model depending on the current circumstances.

With the introduction of e-commerce and e-business in the practice of firms, corporations and banks, there have been changes in the nature of commercial and financial transactions, relations with partners and customers, the development and implementation of business strategies, as well as in real competition. All these new phenomena in the DOW point to the relevance of theoretical research in such a complex, multifaceted and insufficiently studied area as
determining the role of ICT in the development of international business in the field of production, services, trade and finance.

Under the influence of ICT, all areas of international business are being modified, modified, adapted to new phenomena and relations in the business environment. ICTs significantly change the nature of economic interaction between dispersed economic units. Studies of such transformations are very relevant and timely, since the organizational schemes that arise in business are based on the network principle and on network interaction, which becomes the fundamental principle, and networks are the fundamental tool of the new organizational scheme in international business.

The integration of networks of various types and scales and the construction of a global information and communication system in the form of the Internet have led to the creation of a fundamentally new business environment, to the emergence of purely networked virtual companies, in which modern high-tech management is carried out on the basis of electronic infrastructure.

The rapid growth of digital technologies has a significant impact on the development of the economy and society. The role of individual technologies – artificial intelligence, the Internet of Things, blockchain—is increasing, which leads not only to the transformation of production processes, but also to the emergence of new markets, a radical change in business models (platforms, ecosystem companies, etc.). Solutions based on digital technologies can optimize processes, simplify scientific and technical cooperation, promote the connectivity of territories and the involvement of small enterprises in the value chain. The key resource of the digital economy is intangible assets (information, knowledge, competencies). The digital transformation of the economy is primarily caused by changes in the information and communication technology sector[6].

However, despite the positive nature of the use of ICTs in the economic and economic activities of modern society, there are certain imbalances and contradictions. Thus, the existing socio-economic inequality between developed and developing countries is the cause of the so-called "digital inequality" and, consequently, different opportunities for accelerating economic development, obtaining high-quality information services and improving the well-being of the population. In these circumstances, ICTs in developing countries cannot have such a significant impact on economic systems and economic relations, nor on the development of national and international business, as is the case in developed countries.

4 CONCLUSIONS

Let's draw conclusions. Today, information technologies are used to solve a wide range of tasks in various fields-from industry, fuel and energy complex and agriculture to housing and communal services, medicine and transport. Commercial companies on a global level compete with each other in the speed of creating and implementing electronic services.

As a result, the level and quality of information and communication technologies determine the prospects and sustainability of the business, so they are a priority for the organization's efforts, its tool for achieving success.
Acknowledgments: This paper has been supported by the RUDN University Strategic Academic Leadership Program.

REFERENCES

[1] Dr. Nazrul Islam The Use of Information and Communication Technology (ICT) and Business Management: Contemporary Issues and Challenges
Statistical Analysis of the Adaptability of Street-Stall Economic Policies in the Post-Epidemic Era Based on SPSS 22.0-Taking Zhejiang Province and Shaanxi Province as Examples

Xiayi Pan1,a, Minyu Huang1,b, Shiyu Huang1,c
bangxiayi1@163.coma, 13278822100@163.comb, 1725269951@qq.comc

Hohai University, Beijing University of Chemical Technology, Sun Yat-sen University, China1

Abstract: Informal employment models, such as the street-stall economy, play a crucial role in enriching the diversification of employment forms and economic recovery in the post-epidemic period. Taking Zhejiang Province and Shaanxi Province as examples, this paper used SPSS 22.0 to carry out descriptive statistical analysis on the relevant data of 21 urban areas in the two provinces, and compares the economic indicators of the street-stall economic policies. It is found that all cities have issued relevant policies for employment, while large and medium-sized cities in the east focus more on the organic combination of the street-stall economy and urban cultural ecology.

Keywords: Post-Epidemic Era, Informal Employment, Street-Stall Economic Policy.

1 INTRODUCTION

The street-stall economy refers to a marginalized and informal form of economy in which individual workers earn income by providing legal goods or services in public space. There are two types: fixed vendors and mobile vendors. Under the traditional urban management mode, setting up roadside stalls will bring many problems, such as occupying public space and land, polluting the environment, affecting public safety, and difficulty in supervision. Therefore, city managers hold a “one size fits all” attitude towards the street-stall economy.

However, with the outbreak of COVID-19, we have gradually entered the post-epidemic era. In order to alleviate the unemployment problem caused by COVID-19, during the 20th National People’s Congress and the Chinese People’s Political Consultative Conference, Premier Li Keqiang pointed out that it is necessary to reserve living space for the food stall economy and give the people at the bottom of society a little warmth. Later, during his visit to Shandong on June, he once again pointed out that the “street-stall economy” and “small-store economy” are important sources of employment. This has led to an upsurge of “all people setting up stalls”. With the support and advocacy of the government, street-stall economy has also gained temporary legitimacy, and the living conditions of the people at the bottom of society have been alleviated and improved to a certain extent.

However, the contradiction between the characteristics of the street-stall economy and urban management has not been fundamentally resolved. In order to explore how to solve this
contradiction, through data retrieval and comparative research, this paper finds that the economic policy of the street-stall should be adapted to local conditions, and the economic and informal employment policies of different regions should be differentiated. Through the study of policy differences in different regions, we further find out the influence factors of street-stall economy on regional informal employment.

2 RESEARCH BACKGROUND

Since the reform and opening up, China’s urbanization process has accelerated, and new urban immigrants have become the main source of informal employment. Since 1990, the scale of informal employment in cities and towns in China has increased from 30 million to 176 million in 2010, and the proportion of total urban employment has increased from 17.4% to 50.7% [1].

Without the informal sector, the concept of informal employment cannot be defined. In 1973, the International Labor Organization (ILO) first proposed the concept of the “informal sector” in a report entitled “Employment, Income and Equality”. Kenya’s strategy is to increase productive employment. Informal employment refers to employment in the informal sector. According to the interpretation of China’s Ministry of Human Resources and Social Security, informal employment refers to “employment behavior that employee has not signed a labor contract but has formed an actual labor relationship”. It can be summarized as “economic activities occurring in small-scale production and service units and self-employment in urban areas of developing countries” [8]. The employment of urban vendors is an important component of informal employment. It is estimated that the number of urban vendors accounts for 5.2% of the total employment in the city and 15.9% of the total informal and commercial volume in the city [2].

The existing research on street-stall economy and informal employment mainly focuses on the characteristics of street-stall economy and the existing governance problems. Some scholars are concerned about the survival dilemma of vendors, which extends to the social security of informal employees represented by vendors. Zhou Xiaoyan found in the Survey Report of Beijing Non-public Enterprise Labor Relations that under the employment status of the 349 non-public enterprises surveyed, the employment rate of informal employees accounted for 49.56%, among which, temporary workers account for 95% of regular workers. Most of the informal employees represented by vendors are facing problems such as wages, employment difficulties, and lack of insurance and welfare systems. In the past, there were misunderstandings about the perception and management of the street-stall economy. For example, according to the Marxist point of view, the “street-stall economy” is a backward production situation. They are scattered, small in scale, backward in technology and low in labor productivity, and belong to the small commodity economy [3]. The epidemic has had a profound impact on China’s economy. From the macroeconomic, microeconomic and industrial perspectives, Zhiyan and Luo Changyuan found that the epidemic exerted a huge impact on the consumption and service industries, especially on small and medium-sized enterprises. In the post-epidemic era, some scholars have begun to pay attention to specific measures to promote the sustainable development of the stall economy. For example, An Baoyu believed that the construction of a multi-party linkage mechanism and the establishment and improvement of industry norms and access mechanisms can enable the food stall economy to enter a new era and play a role in expanding domestic
demand and consumption. At present, the research on policies in the post-epidemic period is mostly based on specific provinces and regions for empirical analysis. There are few horizontal comparative studies on street-stall economic policies in large, medium and small-sized cities in the post-epidemic era.

At present, a large number of related studies on informal employment have been carried out at home and abroad. The existing research mainly focuses on the influence factors of informal employment choice and the influence of informal employment on different groups. For example, Liu Yan and Li Yueyun (2007) found that elderly women will be more likely to access informal employment. Zhang Shubo and Cao Xinbang (2017) found based on CGSS data that the household registration system makes invisible identity discrimination exist, which has a 13% impact on the choice of formal employment and informal employment. Marc et al. (2009) found in a study of the employment situation in Latin America that people with low levels of education are more inclined to informal employment. Relevant research on employment policy showed that Japan promotes the informal employment of Japanese women by continuously improving the social security and welfare policies related to informal employment and the employment policies of developed countries, and obtains the experience and enlightenment of China’s employment in promoting the improvement of employment quality [5]. Goldberg and Pavcnik (2003) studied the trade policies of Brazil and Colombia and found that the more liberalized the trading system is, the more likely people are to engage in informal employment. The research methods are mostly questionnaires and in-depth interviews. In the data analysis method, the basic theory is mainly used.

3 THE CONCEPT AND CLASSIFICATION OF INFORMAL EMPLOYMENT

Informal employment originated from the “informal sector” elaborated by the International Labor Organization (ILO) in 1973. At present, the theories on informal employment mainly include dual economic theory, poverty employment theory and neoliberal theory. In 2003, the International Labor Conference further improved the framework of informal employment, but in view of the differences in economic development and social systems, domestic scholars have further defined the concept of informal employment respectively in China. But in short, its core characteristics, that is, “non-agricultural employment in cities and towns, different from traditional fixed employment methods and lack of social security” have been generally recognized by domestic scholars. (Qu Yan, 2018)

3.1 Existing Research Results On Informal Employment

Through the keyword search of “informal employment” and “informal sector”, sorted by the number of citations, we find that the existing research on informal employment focuses on the causes of informal employment, the impact of informal employment and related policy measures. As policies and governance measures have been summarized above, they will not be repeated here. See the table below for details.
<table>
<thead>
<tr>
<th>Reasons for informal employment</th>
<th>Labor supply</th>
<th>Demographic characteristics</th>
<th>Age</th>
<th>Jin Yihong believed that the female labor force has a tendency of younger age in the field of formal employment.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Household Registration</td>
<td></td>
<td></td>
<td>Zhang Guoying held it that the dual economic structure and household registration system are important factors that make it difficult for the rural population to enter the formal employment sector.</td>
</tr>
<tr>
<td></td>
<td>Population</td>
<td></td>
<td></td>
<td>Li Qiang and Yang Zhaung estimated that there are more than 0.5 billion people in informal employment, including agricultural production, and the pressure of surplus labor needs to rely on informal employment.</td>
</tr>
<tr>
<td></td>
<td>Human capital characteristics</td>
<td>Technical level</td>
<td></td>
<td>Wan Xiang Dong deemed that the low level of education and labor of workers will be hindered to enter the formal employment sector, thus turning to the labor-intensive formal employment sector.</td>
</tr>
<tr>
<td>Economic environment</td>
<td>Macroeconomic environment</td>
<td>Economic system structure</td>
<td></td>
<td>Hu Angang and Yang Yunxin considered that the structural adjustment of economy leads to the that of employment, and the denationalization of ownership leads to the deregulation of urban employment.</td>
</tr>
<tr>
<td></td>
<td>Income difference</td>
<td>Engel coefficient difference</td>
<td></td>
<td>Through field research, Zhang Huachu found that most of the consumption of laid-off workers in urban informal employment is used for living consumption, and the Engel is 12.6% higher than that of on-the-job employees.</td>
</tr>
<tr>
<td>Administrative reasons</td>
<td>Favorable policies</td>
<td>Government investment promotion</td>
<td></td>
<td>Through meso-level research on informal employment, Huang Zongzhi found that local government intervention, such as investment promotion, has a positive effect on the generation of informal employment groups.</td>
</tr>
<tr>
<td>The impact of informal employment</td>
<td>Economic growth</td>
<td>Driving regional economic development</td>
<td></td>
<td>Yang Manzhen summarized the role of the street-stall economy on the urban economy, and found that the street-stall economy can drive consumption and production by accelerating the flow of small commodities, and promote the development of the national economy.</td>
</tr>
<tr>
<td></td>
<td>Positive impact</td>
<td>Poverty alleviation in rural areas</td>
<td></td>
<td>Through the analysis of the advantages and disadvantages of the development of formal employment, Zhang Huachu found that the remittance of migrant workers to rural areas has greatly improved the income level of rural families.</td>
</tr>
</tbody>
</table>
Informal employment is a part of urban employment and has important positive value and significance. In terms of social effects, it is conducive to absorbing surplus labor and avoiding urban riots caused by unemployment; in terms of economic effects, it is specifically manifested in the pulling effect on GDP. For instance, through interpolation and estimation of tourism multipliers, Guo Wei et al. verified the functional role of the tourism informal sector in stimulating GDP and employment (Guo Wei et al., 2014). The negative impacts can be guided and controlled through policies. If the informal employment is deemed as “illegal employment” from a conventional view, and the means of combat or suppression were taken accordingly (Li Qiang and Tang Zhuang, 2002), this might result in the low-efficient governance and the aggravation of the street poverty. Due to the diversity of informal sector development (Chen, 2006) caused by different levels of economic development, demographic structure and human capital factors in different regions, the regional governance policies will be adopted. At the same time, on account of the important differences in the nature and causes of the informal sector (Grxhani, 2004), the effect of policy governance is closely related to the scale and structure of the city (Huang, Xue, 2011).

### 3.2 Informal Employment and the Street-Stall Economy

The street-stall economy is one of the informal economic forms that emerged in the post-epidemic period (Liu, 2021). It has the characteristics of low entry barriers, less control, and small scale of the informal economy, and it is different from other informal economic forms. It can be defined as an economic form that obtains a source of income through setting up roadside booths (Yang, 2015). From a policy point of view, the two sessions of the National People’s Congress in 2020 will “stabilize employment and protect people’s livelihood”, so that the street-stall economy has a certain legitimacy. In terms of its connotation, the traditional street-stall...
economy is a marginal urban form. After the rapid promotion of e-commerce model in recent years, the street-stall economy has also achieved transformation and upgrading.

The street-stall economy is conducive to promoting the recovery and development of the national economy in the post-epidemic period, improving the employment situation, alleviating poverty and reducing costs (Zhang, 2002). Owing to the three characteristics of low threshold for entrepreneurship, low risk of failure and low commodity price, the street-stall has unique advantages. In the future, the street-stall economy can be used as a transitional form of shop economy [9].

3.3 Street-Stall Economy and Corresponding Policies

As a form of informal economy, street-stall economy is of great positive significance to economic development and social stability. However, due to its difficulties in governance and low economic efficiency, traditional governance thinking is easy to “illegalize” it. After the epidemic, the focus of social governance tended to “stabilize employment and protect people’s livelihood”, and the social benefits contained in the street-stall economy were gradually tapped. The corresponding economic policy of street-stall tends to act as a “social stabilizer”, reflecting the integration of social and economic benefits (Sun, 2009).

The stall economy is one of the informal economy forms. Active street-stall economic policies are easy to play a positive role in the informal economy, which is beneficial to reducing the employment threshold, alleviating the pressure on the lives of the people at the bottom of society, maintaining social stability, and stimulating economic growth through the internal cycle of consumption, etc. By virtue of the diversity of the informal economy (Chen,2006) and geographical differences (Grxhani,2004), there are differences in the economic policies of large, medium and small cities in the post-epidemic period, as shown in the following table:

<table>
<thead>
<tr>
<th>Table 2 The contribution of influencing factors of policy differences in large, medium and small cities</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Variables</strong></td>
</tr>
<tr>
<td>Eastern</td>
</tr>
<tr>
<td>City size</td>
</tr>
<tr>
<td>Policies to promote informal employment</td>
</tr>
<tr>
<td>Policies to encourage the development of street-stall economy</td>
</tr>
<tr>
<td>Topic</td>
</tr>
<tr>
<td>--------------------------------------------</td>
</tr>
<tr>
<td>Income gap</td>
</tr>
<tr>
<td>Gross regional product</td>
</tr>
<tr>
<td>GDP Index (%)</td>
</tr>
<tr>
<td>Number of registered population (10,000)</td>
</tr>
<tr>
<td>Number of permanent population (10,000)</td>
</tr>
</tbody>
</table>
The formulation of relevant policies directly affects the development of informal employment. The proportion of urban population is used to reflect the urbanization rate, including the agricultural population and the floating population in the urban area. Hu Angang and Zhao Li held that the economic value created by the transfer of rural labor affects the extent to which real GDP is underestimated. Here, fiscal expenditure is used to reflect the government’s management functions. Jing Sijiang believed that fiscal expenditure reflects the degree of government support for public services, affects the environment and quality of informal employment, and realizes the transition of the informal sector to decent employment. Education level is used to represent human capital. Wu Yuanwu and Cai Fang thought that the education level determines the competitiveness of workers and has a negative contribution to entering informal jobs. This in turn affects the scale of urban non-employment, including the street-stall economy, and thus influences the formulation of the street-stall economy policy.

Note: This data is selected for the data in 2019, except for fiscal expenditure. That is, because the data in 2019 are the latest data available for most indicators. For the sake of unification, the data under the year with the widest coverage are selected to facilitate horizontal comparative research.

4 ANALYSIS AND DISCUSSION

In this paper, SPSS 22.0 is used to carry out descriptive statistical analysis on the relevant data of 21 urban areas in Zhejiang and Shaanxi Provinces, and the results are shown in Table 2. Tables 3 and 4 show the descriptive statistics of 11 cities in Zhejiang Province and 10 cities in Shaanxi Province.

As can be seen from Tables 2 to 4, for Zhejiang and Shaanxi Provinces, most regions have promulgated policies to promote informal employment in the post-epidemic era (average value of $p_{informality} = 0.81$), and some regional governments have promulgated policies to encourage the development of the vendor economy (average value of $p_{vendor} = 0.57$).
According to the statistics released by the two provinces in 2019, there is a big difference in the number of on-the-job employees and the proportion of urban population between the eastern and western regions. Hangzhou City, Zhejiang Province has the largest number of on-the-job employees, with 4.2546 million, and the proportion of urban population is the largest, which is 78.5%. Yan’an City, Shaanxi Province has the least number of on-the-job employees, with 34,800, and the proportion of urban population in Weinan City, Shaanxi Province is the least of 48.5%. In addition, there are great differences in GDP and fiscal expenditure between regions.

### Table 3 Overall descriptive statistical analysis (N = 21)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Average</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>east</td>
<td>0</td>
<td>1</td>
<td>0.52</td>
<td>0.51</td>
</tr>
<tr>
<td>size</td>
<td>1</td>
<td>3</td>
<td>1.48</td>
<td>0.68</td>
</tr>
<tr>
<td>p informality</td>
<td>0</td>
<td>1</td>
<td>0.81</td>
<td>0.40</td>
</tr>
<tr>
<td>p vendor</td>
<td>0</td>
<td>1</td>
<td>0.57</td>
<td>0.51</td>
</tr>
<tr>
<td>tincome</td>
<td>25503.00</td>
<td>66068.00</td>
<td>46863.90</td>
<td>14390.70</td>
</tr>
<tr>
<td>tincress</td>
<td>7.80</td>
<td>9.10</td>
<td>8.34</td>
<td>0.31</td>
</tr>
<tr>
<td>rincome</td>
<td>10025.00</td>
<td>37413.00</td>
<td>22505.29</td>
<td>11065.18</td>
</tr>
<tr>
<td>rincress</td>
<td>8.70</td>
<td>10.20</td>
<td>9.59</td>
<td>0.47</td>
</tr>
<tr>
<td>GDP</td>
<td>354.72</td>
<td>15418.80</td>
<td>4175.65</td>
<td>3930.91</td>
</tr>
<tr>
<td>GDPIndex</td>
<td>1.90</td>
<td>109.20</td>
<td>58.87</td>
<td>52.01</td>
</tr>
<tr>
<td>njob</td>
<td>3.48</td>
<td>425.46</td>
<td>109.58</td>
<td>115.83</td>
</tr>
<tr>
<td>salary</td>
<td>56628.00</td>
<td>101069.00</td>
<td>73405.43</td>
<td>11174.96</td>
</tr>
<tr>
<td>household</td>
<td>79.26</td>
<td>1295.29</td>
<td>449.70</td>
<td>279.07</td>
</tr>
<tr>
<td>resident</td>
<td>69.83</td>
<td>1036.00</td>
<td>452.50</td>
<td>289.62</td>
</tr>
<tr>
<td>non-household</td>
<td>-274.94</td>
<td>251.24</td>
<td>2.81</td>
<td>113.27</td>
</tr>
<tr>
<td>urbanization</td>
<td>48.50</td>
<td>78.50</td>
<td>62.10</td>
<td>9.03</td>
</tr>
<tr>
<td>expense</td>
<td>47.05</td>
<td>2069.70</td>
<td>537.98</td>
<td>526.45</td>
</tr>
</tbody>
</table>

### Table 4 Descriptive statistical analysis of Zhejiang Province (N = 11)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average</th>
<th>95% Confidence interval for the average</th>
<th>p</th>
<th>Average of the integer</th>
<th>Variance</th>
<th>Standard variation</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Standard error</th>
<th>Lower limit</th>
<th>Upper limit</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>tincome</td>
<td>59212.09</td>
<td>54796.13 6 63269.46 1 59541.24 3 6603 87 000</td>
<td>63222 9601 91</td>
<td>6574417</td>
<td>46437.00 0</td>
<td>66099 000</td>
<td>19831 00</td>
<td>4907 00</td>
<td>0.661</td>
<td>1.279</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard error</td>
<td>1982.261</td>
<td>0.661 1.279</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tincress</td>
<td>8.373</td>
<td>8.099 8.647 8.364 8.5 00</td>
<td>0.166</td>
<td>0.408</td>
<td>7.800</td>
<td>9.100</td>
<td>1.300</td>
<td>0.700</td>
<td>0.262</td>
<td>0.782</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard error</td>
<td>0.123</td>
<td>0.661 1.279</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rincome</td>
<td>32118.818</td>
<td>28479.05 3 35750.70 3 32390.68 7 346 03 000</td>
<td>29312 2301 64</td>
<td>5493523</td>
<td>21931.00 0</td>
<td>37413 000</td>
<td>15482 00</td>
<td>6521.00</td>
<td>0.627</td>
<td>0.808</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard error</td>
<td>1634.04</td>
<td>0.661 1.279</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rincress</td>
<td>9.327</td>
<td>9.012 9.643 9.315 9.2 00</td>
<td>0.226</td>
<td>0.409</td>
<td>0.700</td>
<td>10.10</td>
<td>0</td>
<td>1.400</td>
<td>0.900</td>
<td>0.130</td>
<td>1.177</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard error</td>
<td>0.141</td>
<td>0.661 1.279</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDP</td>
<td>5072.844</td>
<td>2682.657 8652.490 5370.046</td>
<td>511 40 50</td>
<td>19831 201 2 01</td>
<td>4635 674</td>
<td>1371 680</td>
<td>15418 00</td>
<td>14047 20</td>
<td>0</td>
<td>0122.600</td>
<td>1.313</td>
<td>1.322</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>95% Confidence interval for the average</td>
<td>95% Average of the integer</td>
<td>Median</td>
<td>Variance</td>
<td>Standard deviation</td>
<td>Minimum</td>
<td>Maximum</td>
<td>Range</td>
<td>Inner [lower, upper] range</td>
<td>Skewness</td>
<td>Kurtosis</td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>---------</td>
<td>----------------------------------------</td>
<td>---------------------------</td>
<td>--------</td>
<td>----------</td>
<td>-------------------</td>
<td>---------</td>
<td>---------</td>
<td>-------</td>
<td>-----------------------------</td>
<td>-----------</td>
<td>----------</td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>33.28</td>
<td>30.019, 36.540</td>
<td>32.235, 35.484</td>
<td>31.780</td>
<td>2.471</td>
<td>1140.93</td>
<td>8.579</td>
<td>10.084</td>
<td>1.505</td>
<td>-0.083, 1.061</td>
<td>0.827</td>
<td>1.334</td>
<td></td>
</tr>
<tr>
<td>Standard error</td>
<td>1641</td>
<td>1060.04</td>
<td>11000.75</td>
<td>11001.61</td>
<td>11879.00</td>
<td>1228.65</td>
<td>1486.25</td>
<td>1520.35</td>
<td>16347.00</td>
<td>4090.750</td>
<td>0.827</td>
<td>1.334</td>
<td></td>
</tr>
<tr>
<td>Standard error</td>
<td>8.30</td>
<td>8.184</td>
<td>8.417</td>
<td>8.380</td>
<td>0.027</td>
<td>0.165</td>
<td>0.810</td>
<td>0.6400</td>
<td>0.5400</td>
<td>-0.250, 0.574</td>
<td>0.687</td>
<td>1.334</td>
<td></td>
</tr>
<tr>
<td>Standard error</td>
<td>0.05</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.687</td>
<td>1.334</td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td>1193</td>
<td>1086.04</td>
<td>11000.75</td>
<td>11001.61</td>
<td>11879.00</td>
<td>1228.65</td>
<td>1486.25</td>
<td>1520.35</td>
<td>16347.00</td>
<td>4090.750</td>
<td>0.827</td>
<td>1.334</td>
<td></td>
</tr>
<tr>
<td>Standard error</td>
<td>47.54</td>
<td>40.64</td>
<td>54.41</td>
<td>42.980</td>
<td>0.046</td>
<td>0.253</td>
<td>4.400</td>
<td>6.0800</td>
<td>3.425</td>
<td>-0.295, 0.098</td>
<td>0.827</td>
<td>1.334</td>
<td></td>
</tr>
<tr>
<td>Standard error</td>
<td>0.08</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.687</td>
<td>1.334</td>
<td></td>
</tr>
<tr>
<td>GDP</td>
<td>2529</td>
<td>673.90</td>
<td>4184.205</td>
<td>2272.511</td>
<td>1746.18</td>
<td>2025.83</td>
<td>320.83</td>
<td>354.35</td>
<td>9321.20</td>
<td>1688.080</td>
<td>2.372</td>
<td>6.134</td>
<td></td>
</tr>
</tbody>
</table>

Table 5 Descriptive statistical analysis of Shaanxi Province (N = 10)
<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean (95% Confidence Interval)</th>
<th>Median</th>
<th>Variance</th>
<th>Standard Deviation</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>820.98 (810.60 - 831.36)</td>
<td>812.75</td>
<td>6.400</td>
<td>1.900</td>
<td>7.000</td>
<td>6.000</td>
<td>6.000</td>
<td>-0.932</td>
<td>0.005</td>
</tr>
<tr>
<td>GDP per person</td>
<td>5.65 (0.45 - 7.56)</td>
<td>4.290</td>
<td>5.733</td>
<td>3.614</td>
<td>1.901</td>
<td>1.900</td>
<td>1.900</td>
<td>0.687</td>
<td>1.334</td>
</tr>
<tr>
<td>Number of people</td>
<td>44.5 (7.77 - 82.02)</td>
<td>7.134</td>
<td>38.922</td>
<td>54.978</td>
<td>27.126</td>
<td>6.665</td>
<td>52.342</td>
<td>1.874</td>
<td>7.919</td>
</tr>
<tr>
<td>Standard error</td>
<td>16.5 (5.52 - 52.00)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.687</td>
<td>1.334</td>
</tr>
<tr>
<td>Salary</td>
<td>8707 (6.50 - 899)</td>
<td>60352</td>
<td>74496.30</td>
<td>66682.11</td>
<td>63995.5</td>
<td>505</td>
<td>10074.96</td>
<td>7</td>
<td>56625.00</td>
</tr>
<tr>
<td>Standard error</td>
<td>16.5 (5.52 - 52.00)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.687</td>
<td>1.334</td>
</tr>
<tr>
<td>Housing units</td>
<td>17.17 (100.24 - 717)</td>
<td>675.903</td>
<td>400.322</td>
<td>378.380</td>
<td>1036</td>
<td>46.71</td>
<td>9</td>
<td>331.129</td>
<td>70.28</td>
</tr>
<tr>
<td>Standard error</td>
<td>104.72</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.687</td>
<td>1.334</td>
</tr>
<tr>
<td>Resident household</td>
<td>7.65 (257 - 182.20)</td>
<td>548.310</td>
<td>345.276</td>
<td>326.670</td>
<td>6548</td>
<td>0.055</td>
<td>2</td>
<td>255.891</td>
<td>60.83</td>
</tr>
<tr>
<td>Standard error</td>
<td>80.92</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.687</td>
<td>1.334</td>
</tr>
<tr>
<td>Non-household</td>
<td>71.86 (60 - 120.78)</td>
<td>-14.940</td>
<td>-64.270</td>
<td>-30.230</td>
<td>8311</td>
<td>099</td>
<td>79.968</td>
<td>274.94</td>
<td>5400</td>
</tr>
<tr>
<td>Standard error</td>
<td>25.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.687</td>
<td>1.334</td>
</tr>
<tr>
<td>Urbanization</td>
<td>65.2 (15 - 90.943)</td>
<td>50.401</td>
<td>55.060</td>
<td>52.055</td>
<td>64.96</td>
<td>2</td>
<td>8.000</td>
<td>46.90</td>
<td>74.61</td>
</tr>
<tr>
<td>Standard error</td>
<td>2.54</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.687</td>
<td>1.334</td>
</tr>
<tr>
<td>Expenses</td>
<td>191.88 (88 - 97.352)</td>
<td>286.380</td>
<td>180.658</td>
<td>163.040</td>
<td>1789</td>
<td>0.167</td>
<td>132.122</td>
<td>47.05</td>
<td>538.4</td>
</tr>
<tr>
<td>Standard error</td>
<td>41.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.687</td>
<td>1.334</td>
</tr>
</tbody>
</table>

Fig. 1 Stem-and-leaf display of per capita disposable income of rural residents in the east and west
Fig. 2 Stem-and-leaf display of the growth in per capita disposable income of rural residents in the east and west

Fig. 3 Stem-and-leaf display of regional GDP of urban areas in the east and west
Fig. 4 Stem-and-leaf display of average wages of on-the-job employees in the east and west

Fig. 5 Stem-and-leaf display of urbanization rate in the east and west
Fig. 6 Stem-and-leaf display of fiscal expenditure in the east and west

Fig. 7 Stem-and-leaf display of per capita disposable income of urban residents of different urban sizes
Figure 8 Stem-and-leaf display of growth in per capita disposable income of urban residents of different urban sizes

Fig. 9 Stem-and-leaf display of per capita disposable income of rural residents of different urban sizes
Fig. 10 Stem-and-leaf display of growth in per capita disposable income of rural residents of different urban sizes

Fig. 11 Stem-and-leaf display of regional GDP of different urban sizes
Figures 1 to 13 are the stem-and-leaf displays of some relevant indicators of different urban sizes in the east and west. It can be seen from those figures that the economic and political environment of different urban sizes in the east and west is quite different. In view of the differences between the east and the west, it can be found that the per capita disposable income of rural residents in Zhejiang Province is significantly higher than that in Shaanxi Province, but the growth in per capita disposable income of rural residents in Shaanxi Province is significantly higher than that in Zhejiang Province. As can be seen from Figures 3 and 4, the GDP and average wage in the east are higher than those in the west, and the wage in the west is mainly concentrated in the low level within the province, while the wage quartile distribution in the east is relatively even. The urbanization rate in the east is higher than that in the west, but the urbanization rate in Xi’an, Shaanxi Province is significantly higher than that in other cities.
(Figure 5). Figure 6 shows that the fiscal expenditure in the west is generally less, and the fiscal expenditure in the east is higher than that in the west.

For different urban sizes, it can be found that the income level of small cities is relatively lower than that of large and medium-sized cities, but the growth rate is relatively high, while the per capita disposable income gap in large cities is large, which is due to the difference between the east and the west (Hangzhou and Xi’an respectively). Figure 13 shows that there is a positive correlation between urban sizes and fiscal expenditure, and that fiscal expenditure in small and medium-sized cities is relatively low.

In order to implement the spirit of the Implementation Opinions of the General Office of the People’s Government of Zhejiang Province on Further Doing a Good Job in Stabilizing Employment (issued by Zhejiang Office (2020) No. 19) and other documents, many cities in Zhejiang Province have promulgated a series of policies for employment.

Taking Hangzhou as an example, as a big city in the east, relevant government units have issued the Notice on Further Improving the Work of Stabilizing and Ensuring Employment and the Notice on Printing and Distributing the Measures for the Management of Employment Subsidy Funds in Hangzhou, with a view to supporting entrepreneurship to drive employment, supporting multi-channel employment of college graduates, strengthening accurate employment assistance services, optimizing public employment services, and standardizing policy implementation.

Hangzhou has a high level of GDP and residents’ income, and the government has sufficient financial sources to provide reliable economic guarantee for the implementation of the above two related policies through financial expenditure. Operating rules have been formulated around the requirements of a number of policy documents issued by the state and provinces this year, such as epidemic response, skill upgrading, employment of migrant workers and college graduates. Some specific issues in the process of policy implementation are regulated, and the threshold of policy enjoyment is lowered. As a relatively mature big city, Hangzhou has not issued specific policies to encourage the street-stall economy, but through the above-mentioned series of measures to reduce the cost of regularized employment and promote employment.

Ningbo is a medium-sized city in the east. In addition to issuing the implementation rules for further stabilizing employment, the General Office of the Ningbo Municipal People’s Government issued a Notice on the Implementation Plan for Accelerating the Development of the Nighttime Economy in Ningbo (issued by Yongzheng Office (2020) No. 29), which organically combines the development of street-stall economy with the development of urban culture, improves the supply capacity of high-quality products and services for the nighttime economy, and enriches the nighttime economic consumption pattern. Safeguard measures are strengthened to provide support for the main objectives and tasks, while optimizing the spatial layout, developing the consumption field, and implementing characteristic key projects.

Small cities such as Quzhou City, Zhoushan City and Lishui City have also issued relevant notices and taken landing measures in the post-epidemic period to regulate the setting of temporary stalls.

In addition, the General Office of the Xi’an Municipal People’s Government issued a notice on the implementation plan for promoting the revitalization of rural industries in Xi’an. The Xi’an Urban and Comprehensive Law Enforcement Bureau publicized the establishment of temporary
sales points for summer melons and fruits. The Xi’an Urban Management Committee issued a notice on strengthening the establishment and management of temporary roadside stalls (night markets) to promote the development of street-stall economy in the post-epidemic era, and to relax informal employment while conducting reasonable regulations, accordingly, the urban management department is responsible for the management of ensuring the city appearance and environmental order.

5 CONCLUSION

In the post-epidemic period, cities need to take a series of measures to restore economic development and revitalize the market. Taking Zhejiang Province and Shaanxi Province as examples, starting from different urban sizes in the east and west, this paper is centered on the informal employment policies and street-stall economic policies in the post-epidemic era. Research shows that there are similarities and differences in informal employment and street-stall economy of different urban sizes in the east and west. Due to the large discrepancies in the political, economic and cultural levels between the east and the west, there are differences in policy strength and policy implementation guarantees. Among them, Large and medium-sized cities have a high economic level and large fiscal expenditures, which can effectively realize the implementation of policies, promote the diversity of development policies of street-stall economy, and combine the street-stall economy and urban characteristics. While small cities focus more on optimizing management measures.

In addition, this paper also finds that informal employment such as the street-stall economy depends on government policy support. With the relevant documents of the provincial government as the guiding ideology, the characteristic policies and related measures are designated for the municipal cities, which will affect the economic development of the city and even the development of culture and ecological environment. Relevant departments and institutions are organically linked to optimize management and develop beautiful urban and rural areas in China.

REFERENCES

Research on Enterprise Value Evaluation of Electronic Manufacturing Industry

Wensheng Yang\textsuperscript{1,a}, Ruoqi Xin\textsuperscript{1,b,*}

yangwsh@126.com\textsuperscript{a}, *xinrq971023@qq.com\textsuperscript{b}

School of Economics and Management (School of Tourism), Dalian University, No.10, Xuefu Street Jinzhou New District, Dalian116622, China\textsuperscript{1}

Abstract: Based on the operating characteristics of electronic and communication equipment manufacturing enterprises, which are different from other enterprises, the unique value activities of each business are analyzed. On this basis, the formation, change and realization process and characteristics of enterprise value of electronic and communication equipment manufacturing enterprises are clarified, so as to determine the evaluation index of enterprise value. The entropy weight method and AHP method are used to calculate the index weight, and then the grey relational analysis method is used to calculate the comprehensive score and ranking of enterprise value management. Through the analysis of the current management situation of electronic communication equipment manufacturing enterprises, especially the main problems and reasons existing in the management, it is clearly pointed out that the enterprises under the current management system need to update the management concept and improve the management methods and means in their daily operation and management.

Keywords: Value Management, Entropy Weight Method, AHP, Grey Correlation Analysis.

1 INTRODUCTION

Since the reform and opening up, China's electronics and communication equipment manufacturing industry has developed rapidly by relying on its advantages of cheap labor and low-cost land. By undertaking international industrial transfer and contract manufacturing of multinational companies, it has achieved rapid development, occupying a low-end position in the global electronic and communication equipment manufacturing industry chain.

With the disappearance of demographic dividend in China, the weakening of low-cost factor advantage, the global economic downturn, overcapacity, environmental pollution and a series of problems, the sustainable development trend of China's electronic and communication equipment manufacturing industry is not optimistic.

The main problem in the development of electronic and communication equipment manufacturing industry is the lack of core innovation ability. The level of economic development and technological R&D investment objectively limit the development of core innovation capability. Insufficient investment in technology R&D, on the one hand, is related to insufficient attention paid by enterprises, and on the other hand, it is related to the
characteristics of high risk and low conversion rate of R&D investment. In addition, China is in the middle and low end of the global industrial chain, and developed countries often occupy favorable conditions in the core links of R&D and design, gain the profits of the vast majority of products, and set up high technical barriers to China, which is very unfavorable to the development of China's electronic and communication equipment manufacturing industry. The lack of high-quality compound and professional talents also leads to the inability to change the status quo of industry development.

2 LITERATURE REVIEW

Value management is a comprehensive management mode based on enterprise value evaluation and aiming at value growth. Different organizational forms of enterprises have a common goal -- to create value or maximize value. Enterprise theory researchers are devoted to exploring how enterprises should operate to achieve value growth. Based on the theory of value chain, improving the management of enterprise value, analyzing the driving factors of enterprise value, and clarifying the value proposition are the starting points of transformation and development. In this process, it is necessary to take into account the needs of employees, enterprises, customers, supply chain enterprises, industrial chain enterprises and other stakeholders, and provide value-creating products and services for stakeholders. It includes positions focusing on operational efficiency, professional business efficiency, users aiming at customer value, and industrial ecology with platform vision. At the same time, the value creation ability of enterprises also comes from the spontaneous order force within enterprises, and only with the help of scientific and reasonable management structure can value be effectively created and realized, however, the management structure mainly consists of three parts: structural rationality, process rationality and behavioral rationality\(^3\). Some scholars pay attention to the internal management of enterprises and propose that innovative enterprises should choose appropriate value management mode to alleviate contradictions. Employee stock ownership plan, innovative strategic alliance, leading users and process outsourcing and other value management modes can be considered in the list\(^5\). Based on the perspective of value creation, the system can be constructed from three aspects: the formation process of overall value, the key mechanism affecting value, and the evaluation method of enterprise value\(^6\). The focus of value management has different performance in different stages of enterprise development. For example, the formation mechanism of the value of Internet enterprises is significantly different in different stages of the life cycle, such as the start-up stage, growth stage, maturity stage and decline stage. The enterprise value in the start-up stage and growth stage mainly depends on flexible non-financial indicators. The enterprise value in the maturity and decline period is mainly affected by the financial indicators of profitability, and the interaction between financial indicators and non-financial indicators affects the formation of the value of Internet enterprises.

In the context of value management, the enterprise value evaluation system is constructed, and the value chain is used to effectively integrate the overall value and operating value of the enterprise. The evaluation of each important link along the production process is conducive to the discovery of problems existing in the enterprise value management system and the study of optimization strategies, so as to improve the quality of enterprise value management. The study found that service innovation is an important source of value acquisition for manufacturing enterprises, which can be carried out from four aspects: changing business processes, enhancing
customer value, enhancing customer satisfaction and enhancing competitiveness. Traditional enterprise value evaluation lacks consideration of industry factors and less consideration of industrial chain, so it lacks evaluation of long-term investment value of enterprises\(^{(1)}\). Some studies take the value evaluation of listed logistics companies as an example, and carry out comprehensive evaluation combined with the industrial chain of logistics companies, so as to reflect the position and role of the company in the industrial chain. Starting from the overall value of the industry chain, this paper comprehensively considers the service object, market share, dependence on the industry chain, entry and exit difficulty and other factors of logistics companies, takes these factors as important considerations in the value evaluation of logistics companies, analyzes the value point of logistics companies, so as to more accurately evaluate the market value of logistics listed companies.

### 3 SELECTION OF INDICATORS

According to the value chain theory proposed by Michael Porter, there are two forms of value chain. One is the value chain between business activities within an enterprise, which is called enterprise value chain; the other is the value chain between enterprises, which mainly focuses on upstream and downstream enterprises, which is called industry value chain. The index system construction of this paper focuses on two aspects, namely enterprise value chain and industry value chain. At the same time, Porter believes that the value creation of an enterprise, namely the enterprise value chain, is composed of a series of different but interrelated production and operation activities, including basic activities and auxiliary activities. Basic activities include internal logistics, production operations, external logistics, marketing and sales, service, etc., while supporting activities include procurement, technology development, human resource management and corporate infrastructure. Based on the theory of value management, the evaluation index system of value management is constructed. The basic activities of the first-level index include procurement logistics, manufacturing operation and sales. The procurement logistics index in this paper relates to the utilization efficiency of means of transport, which belongs to internal and external logistics in internal and external logistics. Therefore, the procurement logistics index is classified as the basic activity part of the first-level index. Secondary activities in Tier 1 include R&D design, finance, intangible assets and human resources. R&D design belongs to technology development in auxiliary activities, finance belongs to enterprise infrastructure, and human resources belong to human resources management. As a part of the high-tech manufacturing industry, the electronic and communication equipment manufacturing industry is significantly more dependent on patented technology and brand value than the general manufacturing industry. Therefore, the index of intangible assets is added to the first-level index for overall evaluation and measurement. On the basis of the first-level indicators, considering the availability of data and referring to existing research literature, 24 second-level indicators with high frequency are used to construct, as shown in the following table1 and table2:
Table 1: First order index

<table>
<thead>
<tr>
<th>First order index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research and design</td>
</tr>
<tr>
<td>Procurement logistics</td>
</tr>
<tr>
<td>Manufacturing operation</td>
</tr>
<tr>
<td>Sales</td>
</tr>
<tr>
<td>Financial</td>
</tr>
<tr>
<td>Intangible assets</td>
</tr>
<tr>
<td>The human resources</td>
</tr>
<tr>
<td>Supplier (upstream)</td>
</tr>
<tr>
<td>Dealers (downstream)</td>
</tr>
</tbody>
</table>

Table 2: Secondary index

<table>
<thead>
<tr>
<th>Secondary index</th>
</tr>
</thead>
<tbody>
<tr>
<td>R&amp;d investment intensity</td>
</tr>
<tr>
<td>R&amp;d human capital</td>
</tr>
<tr>
<td>Capitalization rate of R&amp;D investment</td>
</tr>
<tr>
<td>Procurement efficiency</td>
</tr>
<tr>
<td>Transport vehicle utilization efficiency</td>
</tr>
<tr>
<td>The number of production employees in the group</td>
</tr>
<tr>
<td>Utilization efficiency of machinery and equipment</td>
</tr>
<tr>
<td>Ratio of labor costs to operating expenses</td>
</tr>
<tr>
<td>Ratio of marketing costs to operating expenses</td>
</tr>
<tr>
<td>Accounts receivable turnover</td>
</tr>
<tr>
<td>Profit margin on sales</td>
</tr>
<tr>
<td>profitability</td>
</tr>
<tr>
<td>Debt paying ability</td>
</tr>
<tr>
<td>Operation ability</td>
</tr>
<tr>
<td>Ratio of intangible assets to total assets</td>
</tr>
<tr>
<td>Intangible assets per share</td>
</tr>
<tr>
<td>Proportion of employees with educational background</td>
</tr>
<tr>
<td>Average annual compensation of employees</td>
</tr>
<tr>
<td>The largest supplier accounted for the total purchase</td>
</tr>
<tr>
<td>The top five suppliers accounted for the total purchase</td>
</tr>
<tr>
<td>Related supplier procurement in the first five accounts for the total procurement</td>
</tr>
<tr>
<td>Largest customer accounts for total sales</td>
</tr>
<tr>
<td>The top five customers accounted for total sales</td>
</tr>
<tr>
<td>The first five related party sales accounted for the total sales</td>
</tr>
</tbody>
</table>

4 INDEX WEIGHT CALCULATION

In the process of comprehensive evaluation, the determination of weight is indeed very important and will have a decisive impact on the final result. The disadvantage of subjective empowerment is that it relies too much on expert opinions, and the result of empowerment is related to the knowledge structure, work experience and preference of the evaluator, with a color of subjective judgment and a certain degree of arbitrariness. When the sample data changes, the weight will also change. According to the statistical rule, with the increase of the sample size, the change of the weight should be smaller and smaller, and eventually tend to a stable value. However, in the actual evaluation process, it is impossible to make the sample data reach a large enough size, and only approximate value can be obtained under the limited samples. It can be seen that both empowerment methods have their own characteristics and shortcomings: subjective empowerment is inheritable and results are stable, but its shortcomings are that the results are highly subjective and limited by researchers’ perspectives and cognition. However, the results of objective weighting are more transparent and objective, but the results are not stable. From the perspective of economic significance, the results do not reflect the relative importance of evaluation indicators, but reflect the degree of dispersion or information content of indicators, which is also a great limitation for the application of economic research. Based on this, the combinatorial weighting method is put forward as formula(1).
This weighting method can just complement each other's weaknesses and take advantage of each other's strengths, so it has good characteristics.

4.1 Entropy weight method

Entropy weight method is an objective evaluation method. Its basic idea is to determine the objective weight according to the variability of the index. In the process of specific use, entropy weight method calculates the entropy weight of each index by using information entropy according to the variation degree of each index, and then corrects the weight of each index by entropy weight, so as to obtain relatively objective index weight. Generally speaking, if the entropy weight of an index is smaller, the greater the variation degree of the index value is, the more information is provided, the greater the role it plays in the comprehensive evaluation, and the greater its weight is. The specific steps are as follows:

4.1.1 Construct judgment matrix

Construct judgment matrix, the formula is as follows:

\[
R' = \begin{bmatrix}
r'_{11} & r'_{12} & \cdots & r'_{1n} \\
r'_{21} & r'_{22} & \cdots & r'_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
r'_{m1} & r'_{m2} & \cdots & r'_{mn}
\end{bmatrix} (i = 1, 2, ..., m; j = 1, 2, ..., n) \tag{2}
\]

4.1.2 Establishing a standardized matrix

Establishing a standardized matrix, the formula is as follows:

\[
R = \begin{bmatrix}
r_{11} & r_{11} & \cdots & r_{11} \\
r_{11} & r_{11} & \cdots & r_{11} \\
\vdots & \vdots & \ddots & \vdots \\
r_{11} & r_{11} & \cdots & r_{11}
\end{bmatrix} (i = 1, 2, ..., m; j = 1, 2, ..., n) \tag{3}
\]

4.1.3 Determine the weight of objective indicators

Determine the weight of objective indicators, the formula is as follows:

\[
h_j = -\frac{1}{\ln m} \sum_{i=1}^{m} a_{ij} \ln a_{ij} (i = 1, 2, ..., m; j = 1, 2, ..., n)
\]

\[
A_{ij} = \frac{r_{ij}}{\sum_{i=1}^{m} r_{ij}} \tag{4}
\]
4.2 Analytic Hierarchy Process

Analytic Hierarchy Process (AHP) is a weight calculation method based on network system theory and multi-objective comprehensive evaluation method. On the basis of in-depth analysis of the nature of complex problems, influencing factors and internal relations. Using less quantitative information, the thinking process of decision making is mathematized. Thus, it provides a simple decision-making method for complex decision-making problems with multiple objectives, multiple criteria or no structure. This method combines qualitative and quantitative analysis, uses the experience of decision makers to judge the relative importance of each indicator, and gives a reasonable weight to each indicator.

4.2.1 Construct judgment matrix

Construct judgment matrix. the formula is as follows:

\[
A = \begin{bmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{n1} & a_{n2} & \cdots & a_{nn}
\end{bmatrix}
\]  

(5)

4.2.2 Check consistency

1. Calculate the elements of each row of the judgment matrix and multiply them to obtain the following result:

\[
M_i = \prod_{j=1}^{n} a_{ij} \quad (i, j = 1, 2, \ldots, n)
\]

\[W^* = \frac{n}{\sqrt{M_i}}\]  

(6)

2. Vector normalization

Vector normalization. the formula is as follows:

\[W_i = W^*/\sum_{i=1}^{n} W^*\]  

(7)

3. Consistency of judgment matrix

To judge matrix consistency, the following formula is used to calculate the value of CI:

\[
AW = \begin{bmatrix}
W_1/W_1 & W_1/W_2 & \cdots & W_1/W_n \\
W_2/W_1 & W_2/W_2 & \cdots & W_2/W_n \\
\vdots & \vdots & \ddots & \vdots \\
W_n/W_1 & W_n/W_2 & \cdots & W_n/W_n
\end{bmatrix} \begin{bmatrix}
W_1 \\
W_2 \\
\vdots \\
W_n
\end{bmatrix} = \lambda \begin{bmatrix}
W_1 \\
W_2 \\
\vdots \\
W_n
\end{bmatrix}
\]
In the process of hierarchical analysis, when the consistency ratio is less than 0.1, it is considered that the degree of inconsistency of the judgment matrix is within the allowable range, and there is satisfactory consistency, and the consistency test is passed.

The data of the AHP in this paper come from the questionnaires filled out by industry experts and professors, and the importance of the indicators is compared pairwise by using the scaling method, and finally the weight of the indicators is obtained. The data rely on the judgment and experience of each expert and are therefore somewhat subjective. Different from the analytic Hierarchy Process (AHP), the data used in the entropy weight method come from 394 A-share listed companies and are manually collated through the 2020 annual reports. It can accurately and realistically reflect the real situation of each company and has certain objectivity. In order to improve the accuracy of index weight calculation, this paper uses a combination of analytic hierarchy process and entropy weight method to calculate index weight.

## 5 SYNTHETIC ANALYSIS

### 5.1 Data standardization

The analysis index system was determined according to the analysis purpose, and the analysis data was collected on valuation objects were set to form the following matrix:

\[
(X'_1, X'_2, ..., X'_m) = \begin{bmatrix}
  x'_1(1) & x'_2(1) & ... & x'_m(1) \\
  x'_1(2) & x'_2(2) & ... & x'_m(2) \\
  \vdots & \vdots & \ddots & \vdots \\
  x'_1(n) & x'_2(n) & ... & x'_m(n)
\end{bmatrix}
\]  

Formula (9)

Due to the different physical meaning of each factor in the system, the dimension of the data is not necessarily the same, which is not convenient for comparison, or it is difficult to get the correct conclusion when comparing. Therefore, in the gray correlation analysis, it is generally necessary to carry out dimensionless processing of data. In this paper, the common equalization method is adopted, that is, to obtain the average value of each index data, and then divide each data by the corresponding average value of the index, as shown in the formula:

\[
X'_n = \frac{x'_n}{x_a}
\]

Formula (10)

The matrix after standardized data processing is obtained:
5.2 Determine sequence of reference

The reference sequence should be an ideal comparison standard. The reference sequence can be constituted by the best and worst values of each index, or other reference values can be selected according to the purpose of evaluation, which are denoted as:

\[ X' = (X'_1, X'_2, \ldots, X'_m) \] (11)

5.3 Determine the maximum and minimum values

Determine the maximum and minimum values, the formula is as follows:

\[
\begin{align*}
\Delta \text{min} &= \min |X'_0(k) - X'_i(k)| \\
\Delta \text{max} &= \max |X'_0(k) - X'_i(k)| 
\end{align*}
\] (13)

5.4 Determine the correlation coefficient and correlation degree

Determine the optimal and the worst correlation coefficient, the formula is as follows:

\[
\delta_i(k) = \frac{\min |X'_0(k) - X_i(k)| + \rho \max_{i=1} |X'_0(k) - X_i(k)|}{|X'_0(k) - X_i(k)| + \rho \max_{i=1} |X'_0(k) - X_i(k)|} \\
(k = 1, 2, \ldots, m)
\] (14)

Determine the optimal and the worst correlation coefficient matrix, the formula is as follows:

\[
(X'_1, X'_2, \ldots, X'_m) = \begin{bmatrix}
X'_1(1) & X'_2(1) & \ldots & X'_m(1) \\
\vdots & \vdots & \ddots & \vdots \\
X'_1(n) & X'_2(n) & \ldots & X'_m(n)
\end{bmatrix}
\] (15)

5.4.3 Relative correlation coefficient and correlation degree

Relative correlation degree = optimal correlation degree (optimal correlation degree + worst correlation degree)

Calculate the relative correlation degree of each company and arrange it in descending order to get the top 15 companies, as shown in the following table 4:
Table 4: Enterprise comprehensive ranking

<table>
<thead>
<tr>
<th>Serial number</th>
<th>Company name</th>
<th>Relative correlation degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Montage Technology Co.,ltd.</td>
<td>0.4524</td>
</tr>
<tr>
<td>2</td>
<td>Hengyu Datacom Aviation Equipment Co.,ltd.</td>
<td>0.4383</td>
</tr>
<tr>
<td>3</td>
<td>Jingjia Microelectronics Co.,Ltd.</td>
<td>0.4373</td>
</tr>
<tr>
<td>4</td>
<td>Sino Wealth Electronic Ltd.</td>
<td>0.4348</td>
</tr>
<tr>
<td>5</td>
<td>Maxscend Microelectronics Co.,ltd.</td>
<td>0.4341</td>
</tr>
<tr>
<td>6</td>
<td>Jushri Technologies, Inc.</td>
<td>0.434</td>
</tr>
<tr>
<td>7</td>
<td>China Spacesat Co.,Ltd.</td>
<td>0.4318</td>
</tr>
<tr>
<td>8</td>
<td>Quantumtect Co.,Ltd.</td>
<td>0.4311</td>
</tr>
<tr>
<td>9</td>
<td>Bestechnic (Shanghai) Co., Ltd.</td>
<td>0.4293</td>
</tr>
<tr>
<td>10</td>
<td>Glarun Technology Co.,ltd.</td>
<td>0.4289</td>
</tr>
<tr>
<td>11</td>
<td>Unistrong Science &amp; Technology Co.,Ltd.</td>
<td>0.4283</td>
</tr>
<tr>
<td>12</td>
<td>All Winner Technology Co.,ltd.</td>
<td>0.4266</td>
</tr>
<tr>
<td>13</td>
<td>SG Micro Corp.</td>
<td>0.4261</td>
</tr>
<tr>
<td>14</td>
<td>Will Semiconductor Co.,ltd.</td>
<td>0.4259</td>
</tr>
<tr>
<td>15</td>
<td>Chengdu Xgimi Technology Co.,ltd.</td>
<td>0.4257</td>
</tr>
</tbody>
</table>

6 CONCLUSION

The value management evaluation index system and value driving factors of electronic and communication equipment manufacturing enterprises restrict and promote each other. With the help of the evaluation index system, value drivers penetrate into the enterprise management, all departments and all employees, giving full play to their role in creating enterprise value. The evaluation system comprehensively and accurately measures and evaluates the value created by the value drivers, giving full play to the value evaluation function of the evaluation system. The comprehensive evaluation results of electronic and communication equipment manufacturing companies are very favorable to the companies in the industry. Through the evaluation of the value management level, we can guide the market to select enterprises with better value management level, establish brands in the industry, and promote the positive and healthy development of the industry. At the same time, it can help enterprises find the gap between themselves and the average level of the industry, understand the existing problems and areas in need of improvement, encourage enterprises to accelerate the implementation of rectification, so as to improve the level of enterprise value management.

REFERENCES


Survey on the Awareness of Basic Drugs among Urban Residents in Jilin Province --Based on Binary Logistic Regression Analysis

Kaiyue Shen1,a, Yanyin Cui1, Jie Chen2, Hui Yue1, Zheng Xie1, Chen Ding1, Jingshuo Liu1, Biao Zhang1,b*

e-mail: 1134707402@qq.com, * Corresponding author: bzhangbiao1018@126.com

School of Health Management, Changchun University of Chinese Medicine Changchun, China1
Changchun Public Security Forensic Center Changchun, China2

Abstract: Objective: To evaluate the awareness of urban residents in Jilin Province of the national essential drug policy. Analyze its influencing factors, to provide a reference for increasing the accessibility and availability of essential drugs. Methods: Questionnaires were used to survey 622 residents within Jilin Province. Results: The level of urban residents' awareness of the national basic drug system in Jilin Province is don't know (13.34%), don't know much (46.78%), know (25.88%), familiar (8.36%), and very familiar (5.63%); the type of medical insurance, whether suffered from chronic diseases, satisfaction with the national basic drug system, attitude toward the free use of basic drugs at the grassroots level, and education were influential factors affecting residents' basic drug use behavior; gender, education, and attitude toward free use of basic drugs at the grassroots level were the influential factors of residents' basic drug awareness. Conclusions: The level of urban residents' awareness of the basic drug policy in Jilin Province needs to be improved. The propaganda of the basic drug system in Jilin Province should be strengthened, and the role of medical personnel in primary medical institutions and grassroots cadres should be brought into play.

Keywords: Urban Residents, Essential Drugs, Cognition.

1 INTRODUCTION

The national basic drug regime is a national drug supply guarantee system built around the basic drug catalog [1]. It is the government's effort to meet public health care needs, ensure people's medicines are safe, effective and reasonable, build out an integrated management system of essential medicines through the rational use of limited medical and health resources[2]. In the report of the 28th World Health Organization Assembly in 1975, the Director-General of the World Health Organization proposed for the first time that each member country could select and purchase essential drugs of reliable quality at a reasonable cost according to the health needs of...
its population. Essential drugs were clearly defined in the "Opinions on the Implementation of
the Establishment of a National Essential Drug System" issued in 2009 that adapt to basic
health care needs, suitable dosage forms, reasonable prices, guaranteed supply, and equitable
public access. This study used a self-made questionnaire to understand the current situation of
urban residents' perceptions of the national essential drug policy in Jilin Province and to analyze
the influencing factors. It provides a scientific basis for improving residents' perceptions of the
essential drug policy, improving their perceived attitudes toward the free use of essential drugs
at the primary level, and increasing their willingness to choose primary care institutions for
medical treatment, and further proposes relevant countermeasures or suggestions for improving
the essential drug policy and increasing the use of essential drugs.

2 OBJECTS AND METHODS

2.1 Object
From May to October 2022, the group randomly distributed questionnaires for the whole Jilin
province, and a total of 1100 questionnaires were distributed and 1018 were collected, of which
622 were valid.

2.2 Methods
This study was conducted using a self-made questionnaire. After the survey was completed, it
was unified by the researcher and then Epidata 3.0 was used to create a database and entered, and
SPSS 25.0 software was used for statistical analysis. Binary logistic regression analysis was used
for the analysis of influencing factors. If not otherwise specified, P is the two-sided
probability and the test level P was set at 0.1.

3 RESULTS

3.1 Basic Demographic Information
A total of 622 residents were surveyed, including 237 males (38.7%) and 395 females (61.3%).
Residents under the age of sixty accounted for 91.31%, 88.26% had education at the
undergraduate level or below, and only 11.74% had a master's degree or above. Most of the
respondents had a monthly income of more than 1,000 yuan, and 89.22% had medical insurance,
of which the most important medical insurance methods were the new agricultural cooperative
and public medical insurance, accounting for 41.00% and 33.60% respectively. Those with
commercial medical insurance only accounted for 20.90%; the results showed that 13.34% of
people did not know the national basic drug policy, 46.78% did not know much about it, 25.88%

knew it, and 8.36% and 5.63% were familiar and very familiar with it respectively. The way to
learn about the basic drug policy is mainly through the Internet accounting for 56.78%, followed
by 40.96% of residents through television. 59.72% of residents prefer to obtain knowledge about
the national basic drug system through short videos app (such as Tiktok and Kuaihou), followed
by Baidu at 43.42%. Although most residents (69.45%) said that their communities (villages and
towns) have not carried out promotional activities for the national essential drug system, the
majority of those who have been promoted still believe that grassroots cadres and medical
personnel in primary care institutions are helpful in promoting the essential drug policy. 44.37% of residents are satisfied with the national essential drug system. 86.50% of residents are supportive of the free use of primary essential 86.50% of the residents are supportive of the free use of basic drugs at the primary level.

3.2 Single-Factor Analysis of Essential Drug Use Behavior and Awareness among Urban Residents in Jilin Province

The results of single-factor analysis showed that gender, education, average monthly income, whether or not they enjoyed the new agricultural cooperative, medical aid and public medical care, whether or not they suffered from chronic diseases, their knowledge of the national essential drug system, whether or not village grassroots cadres (community cadres) were helpful in promoting the essential drug policy, their satisfaction with the national essential drug system, and their attitude toward the free use of essential drugs at the grassroots level on the use of essential drugs by urban residents in Jilin Province the differences were statistically significant (P<0.1) when comparing between groups in terms of behavior. There was statistical heterogeneity in the differences between gender, age, education, whether they enjoyed public medical care, whether their communities (villages and towns) had carried out activities to promote the national essential drug system, whether village grassroots cadres (community cadres) were helpful in the promotion of the essential drug policy, and attitudes toward the free use of essential drugs at the grassroots level on the awareness of essential drugs among urban residents in Jilin Province (P<0.1).

<p>| TABLE 1. UNIVARIATE ANALYSIS OF BASIC DRUG USE BEHAVIOR AND AWARENESS AMONG URBAN RESIDENTS IN JILIN PROVINCE |
|--------------------------------------------------|--------------------------------------------------|-----------------|-----------------|-----------------|
| Features                                          | Essential drug use behavior X² P                  | Awareness of essential drugs X² P                  |
|   Good (n=374)  |   Poor (n=248)  |   Good (n=306)  |   Poor (n=316)  |
| Gender                                            |                                                   |                                                   |
| Male                                              |                                                   |                                                   |
| 148 79                                            | 125 102                                          | 3.83 2                                            | 4.928 0.026     |
| Female                                            |                                                   |                                                   |
| 226 169                                           | 181 214                                          |                                                   |               |
| Age                                               |                                                   |                                                   |
| 25 years old and below                            |                                                   |                                                   |
| 100 83                                            | 73 110                                          | 4.18 4                                            | 17.63 0.001     |
| 26-45 years old                                  |                                                   |                                                   |
| 154 85                                            | 112 127                                          |                                                   |               |
| 46-60 years old                                  |                                                   |                                                   |
| 87 59                                            | 86 60                                            |                                                   |               |
| 61 years old and above                           |                                                   |                                                   |
| 33 21                                            | 35 19                                            |                                                   |               |
| Education                                         |                                                   |                                                   |
| Middle School and below                          |                                                   |                                                   |
| 49 62                                            | 77 34                                          | 19.6 75                                           | 39.93 0.001     |
| High school or junior college                     |                                                   |                                                   |
| 68 44                                            | 69 43                                            |                                                   |               |
| College and undergraduate                        |                                                   |                                                   |
| 184 116                                          | 126 174                                          |                                                   |               |
| Master and above                                 |                                                   |                                                   |
| 73 26                                            | 34 65                                            |                                                   |               |</p>
<table>
<thead>
<tr>
<th>Average monthly income</th>
<th>Less than 1000RMB</th>
<th>1000-2000RMB</th>
<th>2000-3000RMB</th>
<th>3000RMB and above</th>
<th>Type of medical insurance</th>
<th>New Agricultural Cooperative</th>
<th>Commercial insurance</th>
<th>Medical aid</th>
<th>Publicly funded health care</th>
<th>No medical insurance</th>
<th>Chronic disease or not</th>
<th>Yes</th>
<th>No</th>
<th>Whether to understand the national essential drug system</th>
<th>Very familiar</th>
<th>Familiar</th>
<th>Understand</th>
<th>Heard of it, don't know</th>
<th>Don't know</th>
<th>Whether the community (village or town) in which you are located has carried out activities to promote the national basic drug system</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>63</td>
<td>58</td>
<td>13.1</td>
<td>29</td>
<td>0.004</td>
<td>55</td>
<td>66</td>
<td>6.112</td>
<td>0.106</td>
<td></td>
<td></td>
<td></td>
<td>141</td>
<td>114</td>
<td>4.21</td>
<td>3</td>
<td>0.040</td>
<td>135</td>
<td>196</td>
<td>2.42</td>
<td>0.119</td>
<td></td>
</tr>
</tbody>
</table>
Are village grassroots cadres (community cadres) helpful to you in the promotion of the basic drug policy

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
<th>χ²</th>
<th>p</th>
<th>Not promoted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>150</td>
<td>71</td>
<td>6.61</td>
<td>0.037</td>
<td>127</td>
</tr>
<tr>
<td></td>
<td>79</td>
<td>42</td>
<td>0.07</td>
<td>0.787</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>123</td>
<td>106</td>
<td>5.665</td>
<td>0.059</td>
<td>157</td>
</tr>
</tbody>
</table>

Is the promotion of the basic drug policy helpful to you by the medical staff of primary medical institutions

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
<th>χ²</th>
<th>p</th>
<th>Not promoted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>197</td>
<td>46</td>
<td>14.1</td>
<td>0.001</td>
<td>144</td>
</tr>
<tr>
<td></td>
<td>93</td>
<td>36</td>
<td>4.99</td>
<td>0.026</td>
<td>146</td>
</tr>
<tr>
<td></td>
<td>144</td>
<td>45</td>
<td>1.658</td>
<td>0.436</td>
<td>133</td>
</tr>
</tbody>
</table>

Satisfaction with the national basic drug system

<table>
<thead>
<tr>
<th></th>
<th>Satisfied</th>
<th>General</th>
<th>Dissatisfied</th>
<th>No idea</th>
<th>Attitude toward the free use of basic drugs at the grassroots level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>193</td>
<td>111</td>
<td>7</td>
<td>63</td>
<td>Support</td>
</tr>
<tr>
<td></td>
<td>83</td>
<td>73</td>
<td>15</td>
<td>77</td>
<td>347</td>
</tr>
<tr>
<td></td>
<td>31.7</td>
<td>77</td>
<td>1</td>
<td>163</td>
<td>191</td>
</tr>
<tr>
<td></td>
<td>P &lt; 0.001</td>
<td>0.001</td>
<td>P &lt; 0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>127</td>
<td>101</td>
<td>13</td>
<td>65</td>
<td>Oppose</td>
</tr>
<tr>
<td></td>
<td>149</td>
<td>83</td>
<td>9</td>
<td>75</td>
<td>Don't care</td>
</tr>
<tr>
<td></td>
<td>4.797</td>
<td>0.187</td>
<td>0.19</td>
<td>0.18</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>0.001</td>
<td>16.90</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
</tbody>
</table>

3.3 Binary Logistic Regression Analysis of Essential Drug Use Behavior of Urban Residents in Jilin Province

The basic drug use behavior of urban residents in Jilin Province was converted into a dichotomous variable: good = 0, poor = 1. Binary logistic analysis was conducted with basic drug use behavior of urban residents in Jilin Province as the dependent variable and 11 statistically significant variables in the univariate analysis as the independent variables, and the results
showed that the type of medical insurance enjoyed was medical assistance, the absence of chronic diseases, satisfaction with the national basic drug system degree of general, dissatisfaction, and ignorance, and the attitude of indifference to the free use of basic drugs at the grassroots level were the risk factors leading to the poor basic drug use behavior of urban residents in Jilin Province. High school or secondary school, college and bachelor's degree, master's degree and above are protective factors for good medication use behavior.

**TABLE II.** Binary logistic regression analysis of essential drug use behavior of urban residents in Jilin Province

<table>
<thead>
<tr>
<th>Features</th>
<th>β</th>
<th>S.E.</th>
<th>Wald</th>
<th>P</th>
<th>OR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (reference=Male)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>0.158</td>
<td>0.196</td>
<td>0.643</td>
<td>0.422</td>
<td>1.171</td>
<td>0.797</td>
</tr>
<tr>
<td>Education (reference = Middle School and below)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High school or secondary school</td>
<td>-0.638</td>
<td>0.306</td>
<td>4.332</td>
<td>0.037</td>
<td>0.529</td>
<td>0.29</td>
</tr>
<tr>
<td>College and undergraduate</td>
<td>-0.682</td>
<td>0.266</td>
<td>6.556</td>
<td>0.01</td>
<td>0.506</td>
<td>0.3</td>
</tr>
<tr>
<td>Master's degree and above</td>
<td>-1.102</td>
<td>0.375</td>
<td>8.64</td>
<td>0.003</td>
<td>0.332</td>
<td>0.159</td>
</tr>
<tr>
<td>Average monthly income (reference = less than 1000 yuan)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1000-2000 RMB</td>
<td>0.161</td>
<td>0.302</td>
<td>0.283</td>
<td>0.595</td>
<td>1.174</td>
<td>0.65</td>
</tr>
<tr>
<td>2000-3000RMB</td>
<td>-0.387</td>
<td>0.298</td>
<td>1.677</td>
<td>0.195</td>
<td>0.679</td>
<td>0.379</td>
</tr>
<tr>
<td>3000RMB and above</td>
<td>-0.257</td>
<td>0.263</td>
<td>0.949</td>
<td>0.33</td>
<td>0.774</td>
<td>0.462</td>
</tr>
<tr>
<td>Type of medical insurance enjoyed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Agricultural Cooperative</td>
<td>0.042</td>
<td>0.227</td>
<td>0.034</td>
<td>0.853</td>
<td>1.043</td>
<td>0.668</td>
</tr>
<tr>
<td>Medical aid</td>
<td>0.689</td>
<td>0.28</td>
<td>6.069</td>
<td>0.014</td>
<td>1.992</td>
<td>1.151</td>
</tr>
<tr>
<td>Study Question</td>
<td>Mean (SD) 1</td>
<td>Mean (SD) 2</td>
<td>Mean (SD) 3</td>
<td>Mean (SD) 4</td>
<td>Mean (SD) 5</td>
<td>Mean (SD) 6</td>
</tr>
<tr>
<td>--------------------------------------------------------------------------------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Publicly funded medical care</td>
<td>-0.371</td>
<td>0.247</td>
<td>2.244</td>
<td>0.134</td>
<td>0.69</td>
<td>0.425</td>
</tr>
<tr>
<td>Whether suffering from chronic diseases (reference = yes)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>0.783</td>
<td>0.242</td>
<td>10.505</td>
<td>0.001</td>
<td>2.188</td>
<td>1.363</td>
</tr>
<tr>
<td>Whether to understand the national essential drug system (reference = very familiar)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Familiar</td>
<td>0.508</td>
<td>0.523</td>
<td>0.943</td>
<td>0.331</td>
<td>1.662</td>
<td>0.596</td>
</tr>
<tr>
<td>Understood</td>
<td>0.352</td>
<td>0.454</td>
<td>0.601</td>
<td>0.438</td>
<td>1.422</td>
<td>0.584</td>
</tr>
<tr>
<td>Heard of it, don't know</td>
<td>0.282</td>
<td>0.444</td>
<td>0.403</td>
<td>0.526</td>
<td>1.326</td>
<td>0.555</td>
</tr>
<tr>
<td>Don't know</td>
<td>0.532</td>
<td>0.501</td>
<td>1.125</td>
<td>0.289</td>
<td>1.702</td>
<td>0.637</td>
</tr>
<tr>
<td>Are village level grassroots cadres (community cadres) helpful to you in the promotion of the basic drug policy (reference = yes)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>-0.248</td>
<td>0.314</td>
<td>0.623</td>
<td>0.43</td>
<td>0.78</td>
<td>0.422</td>
</tr>
<tr>
<td>Not promoted</td>
<td>-0.173</td>
<td>0.313</td>
<td>0.308</td>
<td>0.579</td>
<td>0.841</td>
<td>0.456</td>
</tr>
<tr>
<td>Are the medical staff of primary medical institutions helpful to you in promoting the basic drug policy (reference = yes)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>0.38</td>
<td>0.337</td>
<td>1.27</td>
<td>0.26</td>
<td>1.462</td>
<td>0.755</td>
</tr>
<tr>
<td>Not promoted</td>
<td>0.403</td>
<td>0.306</td>
<td>1.74</td>
<td>0.187</td>
<td>1.497</td>
<td>0.822</td>
</tr>
<tr>
<td>Satisfaction with the national essential drug system</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.009</td>
</tr>
</tbody>
</table>
Binary Logistic Regression Analysis of Basic Drug Awareness among Urban Residents in Jilin Province

The basic drug awareness of urban residents in Jilin Province was converted into a dichotomous variable: good = 1, poor = 0. A binary logistic analysis was conducted with the awareness of essential drugs among urban residents in Jilin Province as the dependent variable and the seven statistically significant variables in the univariate analysis as the independent variables, the results showed that women, college and bachelor’s degree, master’s degree and above, were protective factors for residents’ basic drug awareness. 46-60 years old, and attitudes of opposition and indifference to the free use of basic drugs at the grassroots level were risk factors for poor basic drug awareness among urban residents in Jilin Province.

### TABLE III. Binary Logistic Regression Analysis of Urban Residents’ Awareness of Essential Drugs in Jilin Province

<table>
<thead>
<tr>
<th>Features</th>
<th>β</th>
<th>S.E.</th>
<th>Wald</th>
<th>P</th>
<th>OR</th>
<th>95%CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (reference = male)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>0.473</td>
<td>0.18</td>
<td>6.886</td>
<td>0.009</td>
<td>1.604</td>
<td>1.127 - 2.284</td>
</tr>
<tr>
<td>Age (reference = 25 years old and below)</td>
<td></td>
<td></td>
<td>3.034</td>
<td>0.386</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26-45 years old</td>
<td>-0.26</td>
<td>0.235</td>
<td>1.228</td>
<td>0.268</td>
<td>0.771</td>
<td>0.486 - 1.222</td>
</tr>
<tr>
<td>Education (reference = Middle School and below)</td>
<td>13.789</td>
<td>0.003</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>--------</td>
<td>-------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High school or secondary school</td>
<td>0.325</td>
<td>0.298</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>College and undergraduate</td>
<td>0.867</td>
<td>0.287</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Master's degree and above</td>
<td>1.192</td>
<td>0.358</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of medical insurance enjoyed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Publicly funded medical care</td>
<td>0.203</td>
<td>0.214</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Has the community (village or town) in which you are located carried out activities to promote the national basic drug system (reference = yes)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>0.192</td>
<td>0.238</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Have village grassroots cadres (community cadres) been helpful to you in promoting the basic drug policy (cf. = yes)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>-0.021</td>
<td>0.278</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not promoted</td>
<td>0.283</td>
<td>0.245</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attitude toward the free use of essential drugs at the grassroots level (reference = support)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oppose</td>
<td>-1.417</td>
<td>0.585</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Don't care</td>
<td>-0.764</td>
<td>0.297</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-0.893</td>
<td>0.356</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4 CONCLUSIONS

4.1 Awareness of National Basic Drug Policy among Urban Residents in Jilin Province

The results of the study found that the percentage of residents who were very familiar, familiar and aware of the basic drug system was 43.44% among the 622 respondents, which is lower than the knowledge rate of 50.7% surveyed by Hu Jinwei et al [12], among residents aged 18 and above, but it also reflects that the promotion of the basic drug system in Jilin Province is not yet in place, especially the low knowledge rate of residents. The main channels for residents to know about the basic drug system are the Internet (56.78%), television (40.96%) and doctors (32.61%). It indicates that internet, television and doctors are still the main channels for residents to learn about the basic drug system. Therefore it can be done by community doctors directly to the residents, or by using the internet and television, especially the internet.

4.2 Strengthen Publicity to Improve Residents’ Basic Drug Use Behavior and Awareness of Essential Drugs

Binary logistic regression analysis showed that satisfaction with the national essential drug system and attitude toward free use of essential drugs at the grassroots level were the influencing factors on residents' essential drug use behavior and attitude toward free use of essential drugs at the grassroots level was an influencing factor on residents' awareness of essential drugs. Residents whose satisfaction level with the national essential drug system is average, dissatisfied, and do not understand, as well as those whose attitude toward the free use of essential drugs at
the grassroots level is indifferent, have poor essential drug use behavior and poor awareness of essential drugs, indicating that this group of residents still lacks a full understanding of the national essential drug system and needs effective state propaganda for disease prevention and treatment, especially the role of the essential drug system in disease treatment should be strengthened so that residents can understand and make use of the policy and reap the benefits. Residents who do not suffer from chronic diseases may not be very concerned about medication application because they are relatively healthy. The poor behavior of residents in the use of essential drugs may also be due to a lack of understanding of the policy, which again needs to be promoted.

4.3 Residents with Low Education Level are the Key Population to Enhance the Awareness and Drug Use Behavior of Essential Drugs

The regression results showed that education and health insurance type were the influencing factors of residents’ essential drug use behavior. Residents who have received education above high school have good basic drug use behavior, and the higher the education, the better, which is closely related to the overall knowledge literacy of residents, educated residents can read the drug use instructions by themselves and understand the knowledge related to drug use to guide the rational and safe use of drugs; such residents who enjoy the type of medical insurance for medical assistance will be poorer compared to those who do not have the financial ability to receive a good education. Education level, age, and attitude toward the free use of essential drugs at the grassroots level are the influencing factors for the awareness of essential drugs among urban residents in Jilin Province, where women and residents with college and bachelor's degree and master's degree and above have good awareness of essential drugs, indicating that women and well-educated residents are more concerned about the national essential drug system and that such people are more concerned about their health. The poor awareness of essential drugs among residents aged 46-60 years may be due to the relatively low proportion of this group of residents who have received a good education.

5 RECOMMENDATIONS

In order to effectively implement and improve the awareness of essential drugs among the resident groups, the following suggestions are made:

Pay attention to the promotion of essential drugs among the resident groups and improve the awareness level of essential drugs among the residents. Strengthen the publicity and promotion of the basic drug system on the Internet, television and other types of media to increase residents’ access to information on basic drugs and raise their awareness of the basic drug policy.

Strengthen the village grassroots cadres (community cadres) to promote basic drugs and visit the residents, if they have a deeper understanding of the basic situation of the residents, they can disseminate the information related to basic drugs to the residents in a targeted and timely and accurate manner according to the situation of different groups of people.

Strengthen the training of medical personnel in medical institutions on basic drug policies and improve the service quality and treatment level of primary medical personnel so that they can provide correct and reasonable guidance to residents in the use of basic drugs.
Strengthen training related to basic drug policies for village grassroots cadres (community cadres) and instruct village grassroots cadres on how to accurately disseminate basic drug policies to residents.

Improve the basic drug system. Strictly regulate the prescribing behavior of basic drugs by medical personnel in primary care institutions, improve the environment and service quality of primary care institutions, and increase the credibility and satisfaction of residents with primary care institutions, thus improving the accessibility of basic drugs.

Improve the essential drug catalog and further improve the selection mechanism of essential drugs. Increase the coverage rate of essential drugs in primary medical institutions, rationalize the layout of primary medical institutions, improve the convenience of medical treatment for residents, and meet the needs of more residents for medical treatment and basic drug use.

Revise the procurement standards for essential drugs, establish a financial subsidy system, reasonably approve the prices of essential drugs, and strengthen the monitoring of the use of essential drugs. Ensure the rational use of essential drugs so that more residents benefit from the essential drug policy.

Promote the free use of basic drugs at the primary level. The free use of basic drugs at the primary level is accepted by most residents, and the free use of basic drugs at the primary level can promote the decentralized treatment of common diseases in primary care institutions and alleviate the problem of difficult and expensive treatment of diseases.

Strengthen the construction of primary medical institutions, strengthen rural government health subsidies, and narrow the gap between urban and rural primary medical institutions.

REFERENCES


Chinese Patent Analysis for Foreign Object Detection in Wireless Charging

Qianxiu Liu¹, Hongshen Pang²,*, Jingdong Tian², Xinghua Wei², Danhui Song²

*Corresponding author's Email: phs@szu.edu.cn

Graduate School of Library Information and Media Studies, University of Tsukuba, Tsukuba, Japan¹
Shenzhen University, Shenzhen, Guangdong, China²

Abstract: As wireless charging is one of the essential technologies of new energy vehicles, China has been keeping pace with the world in the commercial exploration of this emerging technology. This study employs patent analysis tools to analyze the output of Chinese patents on foreign object detection for wireless charging. Patent bibliometric methods are utilized to investigate the focus, hotspots, and frontiers of patent research in this scientific field, which helps researchers to understand the status and development trends of foreign object detection.

Keywords: Wireless Charging, Foreign Object Detection, Patent Analysis, Chinese Patent.

1 INTRODUCTION

As one of the essential basic technologies of new energy vehicles, China has been keeping pace with the world in the commercial exploration of wireless charging technology [2,5]. Domestic research institutions and universities represented by China Electric Power Research Institute, Chongqing University, Southeast University, Chinese Academy of Sciences, and Harbin Institute of Technology have accumulated more than a decade of experience in wireless charging theory and prototypes [1,6]. In addition, domestic companies, mainly represented by ZTE/ZTE New Energy Vehicles, have opened several commercial demonstration lines in more than a dozen domestic cities, such as Chengdu, Changsha, Zhengzhou, Shenzhen, and so on. In recent years, they have accumulated rich practices for the large-scale deployment of wireless charging systems. Foreign Object Detection (FOD) plays a significant role in commercializing wireless charging technologies because it highly determines the operating safety and reliability of the charging system [7-8].

Based on these applications, the local standards for wireless charging have also taken the lead. For example, led by ZTE/ZTE New Energy Vehicle, Shenzhen wireless charging local standards were officially released in 2015. These standards are the most complete and detailed electric vehicle specifications in the current global system. Furthermore, on April 28, 2020, the Standardization Administration of the People's Republic of China announced the release of four national standards for wireless charging systems for electric vehicles [4].

Through the patent analysis, the technology distribution, research trends, competition situation, and technical hotspots of Chinese patents on FOD are presented clearly, which is helpful to
evaluate the current technological development level in the wireless charging field in China from the objective perspective.

2 METHODS AND MATERIALS

2.1 Description of Data Sources and Analysis Tools

IncoPat is a patent database provider from China with a collection of patents from 120 authorities and is updated daily. Using the IncoPat database, the data collection scope is before the search date in 2018 (the search date is up to July 6, 2018). The data type is "Chinese invention application."

2.2 Search Process and Search Formula Development

The final search formula was determined as follows: ((IPC=((H02J5/00 OR H02J7/00 OR H02J7/02 OR H02J17/00 OR G01V3/10 OR G01V3/11 OR G01V3/12 OR G01V3/14))) OR (TIABC=((wireless or inductive or non-contact or remote)) AND TIABC=((charging or power supply or energy transmission or electrical transmission or energy emission or electrical emission or energy transmission or electrical reception or energy reception or electrical reception or coupling resonance or resonant coupling or electromagnetic induction or electromagnetic resonance or magnetic resonance or energy transfer or WPT))) AND ((IPC=(G01V3/11)) OR (TIABC=((Foreign Object Detection OR Foreign Object Detection OR External Object Detection OR Special Object Detection OR Metal Detection OR Metal Object Detection OR Special Object Detection OR Metal Detection OR FOD)))). Approximately 435 patents were retrieved.

3 PATENT TREND ANALYSIS RESULTS

3.1 Trend Analysis

3.1.1 Annual Change in the Number of Patent Applications

Figure 1 shows the trend of the patent application. Since 2011, the number of relevant patent applications in China has rapidly grown, reaching 83 in 2016. The average annual growth rate (PAGR) was 35.7% between 2011 to 2016 (see table 1 to check the calculation formula).

![Figure 1. Annual change in the number of patent applications](image-url)
### Table 1. Analysis index and calculation formula of patent applications (Self-drawing)

<table>
<thead>
<tr>
<th>Analysis index</th>
<th>Calculation formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>The average annual growth rate of patent applications (PAGR)</td>
<td>PAGR = ((\sqrt[\text{Y}]{\frac{\text{NYE}}{\text{NYS}}} - 1) \times 100%)</td>
</tr>
<tr>
<td>Number of Patent Applications in start-up year (NYS)</td>
<td></td>
</tr>
<tr>
<td>Number of Patent Applications at the end of the years (NYE)</td>
<td></td>
</tr>
<tr>
<td>Number of years between the start-up year to the end of the years (Y)</td>
<td></td>
</tr>
</tbody>
</table>

#### 3.1.2 Annual Changes in the Number of Patent Disclosures

Figure 2 shows the annual changes in the number of patent disclosure. The trend of the disclosure can understand the difference in the number of patent disclosure documents of the analyzed object in each period at the macro level. The number of patent disclosures has increased rapidly since 2012, reaching a maximum of 83 in 2017.

![Figure 2. Annual changes in the number of patent disclosures](image)

#### 3.2 Technical Analysis

##### 3.2.1 Overview Analysis of Main Technology Types

From the perspective of the leading technology types of patents, patent technology applications related to FOD in wireless charging are mainly concentrated in the four categories, namely, G01V (geophysics; gravimetry), H02J (an electrical energy storage system), G01N (testing or analyzing materials), and B60L (electric vehicle power units). Among them, G01V and H02J have the most patents.

##### 3.2.2 Technology Application Trends

Figure 3 shows the distribution and development trend of patent applications in different technology directions of the analyzed objects. The result shows that the number of patent applications in G01V and H02J is overgrowing, with patent applications in G01V entering a period of rapid growth from 2010 and reaching a maximum of 59 in 2016. Patent applications in H02J entered a period of rapid growth from 2012 and earned a maximum of 35 in 2015.
3.3 Technology Disclosure Trend

Figure 4 shows the distribution and development trends of the number of patent disclosures in different technical directions of the analysis object. It turns out that the number of patent disclosures in G01V has increased since 2009, reaching 51 in 2016 and 2017. The number of patent disclosures in H02J has increased since 2013, reaching 35 in 2017.

3.3.1 Technology Distribution by Provinces and Cities in China

Figure 5 shows the quantitative distribution of each technical direction of the analysis object in different provinces. Through comparative analysis, it is possible to figure out the regions where important technical movements are concentrated. The relevant patents are mainly concentrated in Guangdong, Jiangsu, Beijing, Shanghai, Anhui, Zhejiang, and other places.
3.3.2 Technology Distribution by Regions in China

Figure 6 shows the quantitative distribution of each technical direction of the analysis object in different cities and only counts Chinese patents. The cities with the most relevant patent applications are Beijing, Shanghai, Shenzhen, and Chengdu.

3.4 Applicant Analysis

3.4.1 Ranking of Applicants

The result shows that applicants of the enterprise type applied for the most significant number of patents, totaling 335. Figure 7 shows that the applicant with the most significant number of related patents is Sony Corporation (17), followed by Panasonic Intellectual Property Management Co., Ltd. (14) and Mettler Toledo Safety Line Co., Ltd. (10).
3.4.2 Patent Application Trend of Applicants

Figure 8 shows the development trend of the number of patent applications filed by each applicant. The result shows that the number of related patent applications of Seiko Epson Co., Ltd. reached the highest value in 2008 (7), the number of Sony Corporation reached the highest value (9) in 2013, and Liuzhou Guotao Technology Co., Ltd. reached the highest value in 2016 (8).

3.4.3 Patent Disclosure Trends of Applicants

Figure 9 shows the development trend of the patent disclosure volume of each applicant. In 2014, the three applicants, Panasonic Intellectual Property Management Co., Ltd., Liuzhou Guotao Technology Co., Ltd., and Dongguan Huadun Electronic Technology Co., Ltd., all achieved the highest number of patent disclosures, with 8, 8, and 6, respectively.
3.4.4 Patent Value of Applicants

Patent value degree is a comprehensive evaluation index about the value of a patent after analyzing the patent regarding more than 20 parameters in three aspects: technical stability, technical advancement, and protection scope. Therefore, studying the distribution of the value degree score of the applicant's patent can provide a macro understanding of the applicant's patent quality and thus objectively evaluate an applicant's competitive strength in terms of patents. Figure 10 shows the distribution of each applicant's patent value degree score. Different colors represent different patent values. It can be seen from figure 10 that Seiko Epson Corporation has the most patent value of 10 points, with six patents. Panasonic Electric Industrial Co., Ltd. amounted to 4, while Mettler-Toledo Safety Line Co., Ltd. and Sony Corporation have two each. In addition, Sony Corporation has 11 patents with a value of 9 points. In contrast, the only Chinese company with the highest score patent is State Grid Corporation of China, with three 7-point patents.
3.5 Legal and Operational Analysis

3.5.1 Transfer Trend

Figure 11 shows an overall upward trend in the transfer of related patents. This analysis is used to understand the trend of the technology cooperation, transformation, application, and promotion of the analyzed object in different periods and to reflect on the operation and implementation of the technology. In addition, by analyzing the changes in the amount of technology transformation, it is possible to understand the direction and popularity of the results of the subject of analysis in different periods, and then predict the development direction and future market application prospects of the technology.

![Figure 11. Transfer Trend](image)

3.5.2 Ranking of Assignors

Table 2 shows the ranking of patent holders according to the number of patents that have been transferred. This analysis identifies the technology export activities of each innovation agent and provides a reference for finding technology holders for transfer. In addition, changes in the direction of technological development or market operations can be inferred.

<table>
<thead>
<tr>
<th>Patent holders</th>
<th>No. of patents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panasonic Electric Industry Co., Ltd</td>
<td>4</td>
</tr>
<tr>
<td>POSTECH Co., Ltd</td>
<td>4</td>
</tr>
<tr>
<td>ADT Services, LLC</td>
<td>3</td>
</tr>
<tr>
<td>Sensor Electronics LLC</td>
<td>3</td>
</tr>
<tr>
<td>SPACON Corporation</td>
<td>2</td>
</tr>
<tr>
<td>Jitong International Limited</td>
<td>2</td>
</tr>
<tr>
<td>Bauerbai Proxi GmbH</td>
<td>2</td>
</tr>
<tr>
<td>Wan Jiasheng</td>
<td>1</td>
</tr>
<tr>
<td>Dongguan Tianguan Energy Saving Technology Co., Ltd</td>
<td>1</td>
</tr>
<tr>
<td>Zhonghui Chuangzhi Wireless Power Supply Technology Co., Ltd</td>
<td>1</td>
</tr>
</tbody>
</table>
3.5.3 Ranking of Assignees

Table 3 shows the ranking of patent assignees according to the number of patents that have been transferred. This analysis reveals the technology introduction of each assignee and predicts the direction of its next technology and market deployment.

<table>
<thead>
<tr>
<th>Patent holders</th>
<th>No. of patents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panasonic Electric Industry Co., Ltd</td>
<td>4</td>
</tr>
<tr>
<td>General Electric Hybrid Technology LLC</td>
<td>4</td>
</tr>
<tr>
<td>ADT Services, LLC</td>
<td>3</td>
</tr>
<tr>
<td>Tyco Fire &amp; Safety Limited</td>
<td>3</td>
</tr>
<tr>
<td>POSTECH Co., Ltd</td>
<td>3</td>
</tr>
<tr>
<td>STATE GRID Corporation of China</td>
<td>2</td>
</tr>
<tr>
<td>Tyco Fire &amp; Safety Limited</td>
<td>2</td>
</tr>
<tr>
<td>POSTECH Co., Ltd</td>
<td>2</td>
</tr>
</tbody>
</table>

3.5.4 Composition of Technology Transfer

Table 4 shows the distribution of technology fields in which patents have been assigned. By analyzing and understanding which technology directions are hot spots for promotion, achievement transformation, and patent operation, determining market maturity, and predicting future development trends. The specialized fields with the most significant number of transferred patents are G01V (21) and H02J (16).

<table>
<thead>
<tr>
<th>Technical fields</th>
<th>Patent numbers</th>
<th>Technical fields</th>
<th>Patent numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>G01V</td>
<td>21</td>
<td>H01F</td>
<td>1</td>
</tr>
<tr>
<td>H02J</td>
<td>16</td>
<td>H01M</td>
<td>1</td>
</tr>
<tr>
<td>G01R</td>
<td>4</td>
<td>H02H</td>
<td>1</td>
</tr>
<tr>
<td>G08B</td>
<td>2</td>
<td>H04B</td>
<td>1</td>
</tr>
<tr>
<td>G01D</td>
<td>1</td>
<td>H05B</td>
<td>1</td>
</tr>
<tr>
<td>G01N</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4 CONCLUSION

Since 2010, Chinese patent applications for FOD in wireless charging have shown a rapid growth trend. From the perspective of the main technology types of patents, the G01V and H02J categories have the most patents, and the number of patent applications has increased rapidly. This study helps researchers to understand the status and development trends of scientific research in FOD for wireless charging.
Acknowledgment: This work was supported by The China Postdoctoral Science Foundation (No. 2019M650803, No.2020T130637) and the Guangdong Province Intellectual Property Project of Collaboratively Enhancing Academic Patent Transfer Achievement (No.HT-2022-CJ-ZY-004). Hongshen Pang is the corresponding author.

REFERENCES

Risk Identification of Bulk Commodity Electronic Trading Based on HHM

Xiangping Deng, Xiaoyan Gu*
xiangpingdeng@bistu.edu, *xiaoyangu@bistu.edu

School of Information Management, Beijing Information Science and Technology University, Beijing, 100192, China

Abstract: There are many participants in bulk commodity electronic trading, and the trading modes are diverse. The factors that trigger risks are complex and changeable, and the risks are difficult to identify. This paper establishes a risk identification framework for bulk commodity electronic trading based on HHM, and identifies the risks of bulk commodity electronic trading from the perspective of multi-agent and system. According to the text data, risk factors are extracted based on the content analysis method. Through further statistical analysis, the trading platform, dealers, trading banks and delivery warehouse are derived as the main risk subjects. Combined with the bulk commodity electronic trading mode and business process, the risk factors of bulk commodity electronic trading are systematically extracted to provide support for the risk management of bulk commodity electronic trading.

Keywords: Bulk Commodity, Electronic Trading, HHM, Risk Identification.

1 INTRODUCTION

Bulk commodity electronic trading, also known as online spot trading or spot position trading, is a trading method that uses computer network to organize the same goods in different places, synchronous centralized bidding or one-way bidding, unified matching, unified settlement, and real-time display of price quotations. With the development trend of the bulk commodity electronic trading market, it has evolved from the simple material circulation to trade financing, financial speculation, and other directions, highlighting the characteristics of multiple agents, diverse trading models and complex and volatile risks. The continuous occurrence of risks has seriously affected the normal operation of the bulk commodity electronic trading market and damaged the interests of the trading subjects. Therefore, how to comprehensively and accurately analyse and identify various risks, and how to effectively integrate these information into a systematic and standardized risk identification framework to meet the current and future business development needs of the bulk commodity electronic trading market is particularly important.

Domestic and foreign scholars have identified and analyzed the risks of the bulk commodity electronic trading market from a macro perspective. Zhou et al. 1 summarized and analyzed the current situation and development trend of the supervision and service technology of the bulk commodity trading market. Cao et al. and Fu through their analysis of the commodity
trading market, emphasized that China's commodity market faces problems such as single and fixed model granularity and poor accuracy of market risk early warning \[2-3\]. Wang et al. used LDA topic model to identify risk elements of electronic trading of commodities based on news texts in the commodities domain \[4\]. Shi and Feng established the risk identification and classification framework of bulk commodity electronic trading market from the perspective of trading centers and traders \[5\]. Wang et al. and Feng et al. systematically organize the main risk involved in electronic trading centers of commodities by analysing major risk events \[6-7\]. More risk analysis work usually focuses on a certain type of risk factors, such as price fluctuation risk, supply risk, financial risk, etc. Bakas and Triantafyllou studied the prediction ability of macroeconomic uncertainty on the volatility of agricultural, energy, metal and other commodity markets \[8\]. Hu and Wu studied the relationship between international commodity price volatility and domestic financial market representative index volatility and found that there is a significant two-way spillover effect between China's financial market volatility and international commodity market volatility \[9\].

Domestic and foreign studies often focus on a single type of risk, lacking systematic analysis of market risk from multiple risk subjects and multiple analysis granularity. At this stage, electronic trading market risk has been characterized by cross-region, cross-platform, multi-agent and multi-mode. The market risk analysis needs to conduct effective dynamic granularity scaling according to the changes of regulatory objects and requirements. Based on this, the paper focuses on the subject risk of bulk commodity electronic trading, analyses the information of bulk commodity electronic trading market from multiple perspectives and dimensions in combination with trading mode and business process, introduces the idea and methodology of HHM(Hierarchical Holographic Model), and proposes a risk identification model of bulk commodity electronic trading market based on HHM. Based on the text data analysis of bulk electronic trading risk cases and events, using the content analysis method to iteratively extract risk elements through statistical analysis, this paper establishes a multi-agent bulk commodity electronic trading market risk identification framework system from the perspective of trading platform, dealers, trading banks and trading positions. The multi-agent risk index identification method of bulk commodity electronic trading proposed in this paper can more systematically extract and identify risks, and provide support for risk management of bulk commodity electronic trading.

2 HHM FRAMEWORK FOR BULK COMMODITY ELECTRONIC TRADING RISK

In the process of bulk commodity electronic trading, there are many subjects involved, which is a complex system engineering to identify the multi-agent risk of bulk commodity electronic trading. HHM model can study the system from multiple angles and multi-dimensions, conduct comprehensive and detailed risk identification analysis on complex multi-agent system, and finally form a relatively complete set of risks.

2.1 HHM Model

Hierarchical Holographic Model is a comprehensive systemic thinking model to express the essence and characteristics of a system, and to identify and assess the risk of the whole system
from multiple dimensions and perspectives. The model structure can be refined through continuous cycles and iterations to identify the various risk elements in the system. The traditional mathematical identification models have limited reflection of various aspects within the system, mainly because they are mostly two-dimensional and cannot describe the whole system precisely, completely and clearly. A system contains not only elements, objectives and constraints, but also other factors. The description of risk factors cannot be completed by using only a single system model. HHM model establishes a system risk factor identification framework with a holographic multidimensional perspective, which better solves the problem of risk factor identification in large systems.

2.2 Construction of Transaction Risk HHM Framework

The bulk commodity electronic trading market is greatly affected by the upstream and downstream of the supply chain, international prices, and market fluctuations. There are many uncertain factors in the trading process, and the trading subjects present multiple correlations. There are many factors and complex scenarios that cause market systemic risk. Kaplan and Garrick gave three sets of risk definitions \( R = \{(S_i, L_i, X_i)\}_C \). In the commodity electronic trading market, \( S_i \) refers to the risk subject of commodity electronic trading. \( L_i \) indicates the possibility of the risk of the third risk subject; \( X_i \) indicates the loss vector or consequence caused by the \( i^{th} \) risk subject. \( C \) indicates that all risk entities in the system are included in the risk entity \( S \). The formula proves that risk identification first needs to find the risk subject of the assessed object, and the comprehensive nature of the risk subject determines the accuracy and reliability of the risk assessment results.

In the process of bulk commodity electronic trading, the price of bulk commodities fluctuates frequently and greatly, which enables dealers to use the price law to make speculative profits or hedge, and destroy the forward transaction for the purpose of goods trading; If the applicable laws and regulations are not perfect, or the trading platform operates in violation of regulations, such as unreasonable forced closing operations, the interests of dealers will be seriously damaged. Furthermore, if the third-party depository institution lacks necessary security measures, there is a risk of misappropriation of funds. To sum up, risk subject is contained in all links of electronic commodity trading. By analysing the process of electronic commodity trading market, the subject information of electronic commodity trading in different trading modes can be mined.

The electronic transaction of bulk commodities involves various links from the preparation of goods before the transaction to the delivery and settlement after the transaction is completed. The main links include warehousing, trading, delivery, settlement, financing, etc. At present, the trading process of bulk commodity electronic trading is mainly shown in Figure 1 below.
Through the analysis and research of several representative and large-scale bulk commodity electronic trading models in China, this paper summarizes several trading models mainly used by bulk commodity electronic trading platform, including spot trading model, listing trading model, bidding trading model, and bidding trading model.

Further, by analysing and studying the trading process of the bulk commodity electronic trading market, we can find that there are many interest entities participating in the bulk commodity electronic trading market. In order to screen the main risk subjects of commodity electronic trading, 72 cases related to commodity electronic trading risk in the past 10 years were collected through mainstream search engines, literature databases, commodity trading platforms, legal documents and other channels, and risk subjects were extracted based on risk cases. Among many risk subjects, the electronic trading platform, trading bank, Delivery warehouse and dealer have the highest frequency.

2.3 Risk Identification Process of Bulk Commodity Electronic Trading

The HHM model is used to analyse and identify the risk factors of the bulk commodity electronic trading market from the multi-agent perspective, and the content analysis method, statistical analysis method and other methods are used to analyse and study the data and models of the literature research on the risk of bulk commodity electronic trading, and the analysis results of the multi-agent system risk factors are obtained, which are brought into the summarized bulk trading risk cases in recent decades for comparative verification. The iterative process of risk analysis is mainly shown in Figure 2 below:

Figure 1: Business flow chart of bulk commodity electronic trading market
Based on the analysis of the trading process of the bulk commodity electronic trading market and the mining research of case data, taking the trading platform, dealers, trading banks and delivery warehouse as the main risk subjects and combining the characteristics of the bulk commodity market, the above steps have been used to complete multiple iterative process analysis, and a relatively complete risk indicator system of the bulk commodity electronic trading market has been obtained.

3 RISK FACTORS IDENTIFICATION OF BULK COMMODITY ELECTRONIC TRADING BASED ON HHM

3.1 Trading Platform Risk

Bulk commodity electronic trading platform refers to the e-commerce platform that provides enterprises and individuals with trading, capital settlement, settlement and other related services. During the operation of the trading platform, due to changes in the external legal environment or various participants of the trading platform do not comply with relevant laws and regulations, it may cause negative legal consequences to the trading platform. In terms of trading, settlement, settlement, financing, risk control, internal management, etc., the implementation and modification of the trading platform are not in place due to the formulation of the system. At the same time, due to the internal reasons of the trading platform, there are problems in the information network system supporting the business operation and the technology and measures to ensure the continuous operation of the business,
which may cause losses to the trading platform. Based on the relevant analysis, identify the risk analysis elements of the electronic trading platform as shown in Table 1.

Table 1: Trading platform risk factors analysis.

<table>
<thead>
<tr>
<th>Risk level 1</th>
<th>Risk level 2</th>
<th>Risk screening path</th>
</tr>
</thead>
<tbody>
<tr>
<td>Qualification risk</td>
<td>Business license</td>
<td>Legal person certificate</td>
</tr>
<tr>
<td></td>
<td>Enterprise credit</td>
<td>Legal person certificate</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Platform announcement</td>
</tr>
<tr>
<td>Institutional risk</td>
<td>Inadequate system</td>
<td>Platform announcement</td>
</tr>
<tr>
<td></td>
<td>Inadequate system implementation</td>
<td>Platform data</td>
</tr>
<tr>
<td>Legal risk</td>
<td>Illegal activities</td>
<td>Legal instrument</td>
</tr>
<tr>
<td></td>
<td>Violations</td>
<td>Administrative penalty news</td>
</tr>
<tr>
<td>Operational risk</td>
<td>Technical risk</td>
<td>Fault record</td>
</tr>
<tr>
<td></td>
<td>Operational risk</td>
<td>Platform data</td>
</tr>
</tbody>
</table>

3.2 Dealer Risk

Dealers refer to enterprise users or individual users who participate in trading or other related businesses on the trading platform and receive relevant services. They are the most active factor in the operation of the bulk commodity electronic trading market. The potential risks that dealers may bring to the parties involved in the transaction and the trading platform in the process of their participation in trading activities due to their low credit rating, or their illegal and illegal behaviours. Meanwhile, the dealer fails to perform the transaction and delivery contract due to the sudden change in the price of the traded commodity, problems in its own operation and management, insufficient loan preparation, or the quality of the goods not in line with the agreement, thus damaging the interests of the observant party and causing serious interference to the market. Based on the relevant case analysis, identify the risk analysis elements of dealers as shown in Table 2. For example, in the Shanghai steel price crash, it was precisely because most of the steel trading enterprises in Shanghai fell into huge losses and the capital chain was broken, and a large number of steel traders went bankrupt, leading to the Shanghai Steel Trade Case.

Table 2: Dealer risk factors analysis.

<table>
<thead>
<tr>
<th>Risk level 1</th>
<th>Risk level 2</th>
<th>Risk screening path</th>
</tr>
</thead>
<tbody>
<tr>
<td>Credit risk</td>
<td>Dealer Credit</td>
<td>Enterprise Credit Information System</td>
</tr>
<tr>
<td></td>
<td>Violations</td>
<td>Enterprise Credit Information System</td>
</tr>
<tr>
<td>Transaction performance risk</td>
<td>Non-performance risk</td>
<td>Platform announcement</td>
</tr>
<tr>
<td></td>
<td>Loan payment risk</td>
<td>Platform Data</td>
</tr>
<tr>
<td></td>
<td>Cargo delivery risk</td>
<td>Platform Data</td>
</tr>
</tbody>
</table>
3.3 Trading Bank Risk

The trading bank provides services for the transactions in the bulk commodity electronic trading market but is vulnerable to the impact of information asymmetry risk, the biggest impact of which is the risk of warehouse receipt pledge. Based on the relevant analysis, identify the risk analysis elements of electronic transaction banks as shown in Table 3. For example, Qingdao Decheng Mining Co., Ltd. and several warehousing companies issued warehouse receipts respectively, and then used the loophole of bank information asymmetry to repeatedly pledge to different banks to defraud loans, leading to the occurrence of risks.

<table>
<thead>
<tr>
<th>Risk level 1</th>
<th>Risk level 2</th>
<th>Risk screening path</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information asymmetry risk</td>
<td>Warehouse receipt pledge</td>
<td>data dictionary</td>
</tr>
<tr>
<td></td>
<td>Fake bill</td>
<td>data dictionary</td>
</tr>
<tr>
<td>Default loss rate</td>
<td>Default loss rate</td>
<td>Data dictionary dynamic risk exposure</td>
</tr>
<tr>
<td>Monetary interest rate</td>
<td>Monetary interest rate</td>
<td>China Money Network</td>
</tr>
</tbody>
</table>

3.4 Delivery Warehouse Risk

The delivery warehouse is the storage place of electronic trading commodities selected by the electronic trading center. In recent years, there have been many repeated pledge events of aluminium ingot warehouse receipts in the aluminium spot market. The main risks involved include non-standard warehousing quality inspection, non-standard storage of goods, and internal and external collusion. Based on the correlation analysis, identify the risk analysis elements of trading warehouse as shown in Table 4.

<table>
<thead>
<tr>
<th>Risk level 1</th>
<th>Risk level 2</th>
<th>Risk screening path</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regulatory pledge of goods risk</td>
<td>Non-standard warehousing quality inspection</td>
<td>Data dictionary</td>
</tr>
<tr>
<td></td>
<td>Irregular preservation of goods</td>
<td>Data dictionary</td>
</tr>
<tr>
<td></td>
<td>Internal and external collusion</td>
<td>Data dictionary</td>
</tr>
<tr>
<td>Standardization of warehouse receipt elements</td>
<td>Standardization of warehouse receipt elements</td>
<td>News Corpus</td>
</tr>
<tr>
<td>Perfection of warehousing elements</td>
<td>Perfection of warehousing elements</td>
<td>News Corpus</td>
</tr>
</tbody>
</table>

3.5 Market Risk

Due to changes in the supply and demand relationship of commodities in the market, macro-control, natural disasters and other reasons, the price of relevant commodities fluctuates abnormally. This fluctuation may be transmitted to the price of trading commodities on the trading platform, which may cause the risk of loss or default of a large number of trading users. Meanwhile, when the national and industrial policies change, it may also lead to the risk that the operation of the trading platform will be greatly affected. Based on the correlation analysis, the market risk factors identified are shown in Table 5.
Table 5: Market risk factors analysis.

<table>
<thead>
<tr>
<th>Risk level 1</th>
<th>Risk level 2</th>
<th>Risk screening path</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price fluctuation risk</td>
<td>Supply and demand changes</td>
<td>News Corpus</td>
</tr>
<tr>
<td></td>
<td>Macro-control</td>
<td>News Corpus</td>
</tr>
<tr>
<td></td>
<td>Natural disaster</td>
<td>News Corpus</td>
</tr>
<tr>
<td></td>
<td>Joint operation</td>
<td>News Corpus</td>
</tr>
<tr>
<td></td>
<td>Insider trading</td>
<td>News Corpus</td>
</tr>
<tr>
<td>Policy risk</td>
<td>Policy risk</td>
<td>News Corpus</td>
</tr>
<tr>
<td></td>
<td>Industry Policy Changes</td>
<td>News Corpus</td>
</tr>
</tbody>
</table>

3.6 Bulk Commodity Electronic Trading Market Risk Indicator System

Based on the risk factors of various stages of bulk commodity electronic trading identified from the perspective of the above trading platform, trading customers, trading banks, delivery warehouse and market system, the risk indicator system of bulk commodity electronic trading market is constructed, as shown in Table 6.

Table 6: Bulk commodity trading market risk indicator system.

<table>
<thead>
<tr>
<th>Risk subject</th>
<th>Risk level 1</th>
<th>Risk level 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trading platform P₁</td>
<td>Qualification risk R₁</td>
<td>Business license</td>
</tr>
<tr>
<td></td>
<td>Inadequate system</td>
<td>Inadequate system implementation</td>
</tr>
<tr>
<td></td>
<td>Legal risk R₃</td>
<td>Illegal activities</td>
</tr>
<tr>
<td></td>
<td>Operational risk R₄</td>
<td>Technical risk</td>
</tr>
<tr>
<td>Dealer P₂</td>
<td>Credit risk R₅</td>
<td>Dealer Credit</td>
</tr>
<tr>
<td></td>
<td>Transaction performance risk R₆</td>
<td>Non-performance risk</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Loan payment risk</td>
</tr>
<tr>
<td>Trading bank P₃</td>
<td>Information asymmetry risk R₇</td>
<td>Warehouse receipt pledge</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fake bill</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Default loss rate</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Monetary interest rate</td>
</tr>
<tr>
<td>Delivery warehouse P₄</td>
<td>Regulatory pledge of goods risk R₈</td>
<td>Non-standard warehousing quality inspection</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Irregular preservation of goods</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Internal and external collusion</td>
</tr>
</tbody>
</table>
4 CONCLUSIONS

HHM can display the different characteristics and properties of the system from multiple angles, which is very suitable for identifying the risks of bulk commodity electronic trading. This paper constructs a framework of bulk commodity electronic trading based on HHM; focusing on the characteristics of the trading platform, dealer, trading banks, delivery warehouses and commodity markets, starting from the analysis of the trading process and business model of the commodity electronic trading market, a risk identification model is constructed, and a list of commodity electronic trading risks based on sustainability is obtained. The HHM method overcomes the shortcomings of the traditional risk identification methods in terms of systematicness and integrity, which can comprehensively and systematically identify the risk factors of the bulk commodity electronic trading market. This study provides an important supplement to the systematic study of commodity risk indicators, builds a comprehensive and feasible risk indicator system, provides support for the risk management of commodity electronic trading, and provides index and calculation support for the subsequent quantitative analysis and early warning of commodity electronic trading market risk.

Acknowledgements: This paper is supported by National Natural Science Foundation of China (71701020), National key research and development program (2019YFB1405003).

REFERENCES


The Influence of Asymmetry of Positive Feedback Trade on Expected return: An Empirical Study of China Securities 800

Qinyi Yu
2006070136@pop.zjgsu.edu.cn

School of Finance, Zhejiang Gongshang University, Hangzhou, China

Abstract: Using the measurement and empirical research on the asymmetry of the stock market, this paper confirms risk premium of positive feedback trading asymmetry and its impact on expected returns in China. Through correlation analysis, it is confirmed that the asymmetry factor can represent the trading intensity of retail investors. The A-CAPM model is constructed by incorporating the asymmetry factor into the explanatory framework of stock expected return, and through the significant change of the improved regression intercept term, it is confirmed that the asymmetry factor can better explain the risk premium. Finally, according to the value of the asymmetric factor, the data are divided into four parts, and it is found that the influence of the asymmetric factor on the expected return is opposite to the sign of the factor, and the influence increases with the increase of the absolute value of the factor.

Keywords: Positive Feedback Trading, Asymmetry, A-CAPM Model, Expected Return, Risk Premium

1 INTRODUCTION

Positive feedback trading, also known as "Chasing gains and Selling losses" in the stock market, is an irrational behavior of making trading decisions based on past stock information, which is specifically manifested as buying stocks when they rise and selling stocks when they fall. The asymmetry of the positive feedback trade is reflected in the difference between the intensity of the chasing and the killing.

However, the current Chinese stock market is still not mature market, and the large number of retail investors is a prominent feature of China's stock market. Ma(2016) proposed that retail investors are prone to dependence on policy guidance, and the inadequacy of the existing disclosure mechanism and supervision means leads to a typical herding effect. At the same time, compared with foreign markets, due to the restriction on short selling in Chinese markets, retail investors are more inclined to chase stocks with rising stock prices, rather than immediately sell down stocks, which reflects the characteristics of chasing gains. Zhao et al. (2001), through empirical analysis of China's stock market, confirmed that China's stock market has a more significant disposal effect than that of foreign countries, that is, it is more inclined to sell winners and hold losers, which will further weaken the intensity of the killing and make the chasing more significant.
For institutional investors, under adverse market conditions, their risk control strategy requires them to further reduce risks through short selling, showing the characteristics of killing. As early as in the last century, Sentana and Wadhwani(1992) had found that in the British and American markets dominated by institutional investors, there was a general phenomenon that the killing was stronger than the chasing. Due to the difference in investor structure, the number of individual investors in China's stock market is far more than that of institutions. Wan et al. (2016) found that Chinese stock market has a remarkable feature of chasing gains rather than selling losses by using empirical data.

Therefore, the positive feedback trading asymmetry reflects the trade intensity balance between retail investors and institutional investors. When the trading intensity of one side is significantly stronger than that of the other side, it will show asymmetry.

Traditional capital asset pricing model (CAPM) explains the relationship between expected rate of return and risk assets, and believes that market risk can fully explain expected rate of return. Later, with the development of empirical research, scholars gradually found that there was still some unexplained market risk in the expected rate of return. Such as size effect (Banz, 1981), value effect (Rosenberg et al.,1985), liquidity risk premium (Amihud and Mendelson, 1986), etc. Since then, there have been many scholars (Wang and Zhou, 2002; Yi, 2005; Zhu and Chen, 2021) verified that the expected rate of return was indeed affected by the above factors through empirical research. However, few studies have explored whether retail trading intensity affects expected returns. Retail trading intensity may affect stock price and stock price volatility through influencing investors' investment expectations, and then affect expected return rate. The research objective of this paper is to explore the impact and path of positive feedback trading asymmetry on expected return rate.

In this paper, the actual data of 800 constituent stocks in China Stock market in the past 20 years will be used as samples to measure and measure the asymmetry of positive feedback trading, and the correlation analysis will be conducted between the positive feedback trading asymmetry factor and the retail preference index. After that, the A-CAPM model is constructed to explain the expected returns by testing the risk premium in the Chinese market and introducing the positive feedback trade asymmetry factor. Finally, according to the positive feedback trade asymmetry factor, the stock is divided into classes, and analyzes the specific impact of positive feedback trade asymmetry factor on expected returns and the reasons.

The innovation of this paper lies in:

1. Improve the CAPM model by using the positive feedback trade asymmetry factor, and construct the A-CAPM model, which can better explain the risk premium other than market risk.

2. Use positive feedback trading factors to divide the samples and obtain different impacts of positive feedback trading factors on expected returns under different positive feedback trading states.
2 LITERATURE REVIEW

DeLong and Shleifer(1990) showed that positive feedback trading would lead to sharp price fluctuations. Fang and Meng(2019) believed that herding effect would increase stock price fluctuations caused by positive feedback trading in a short time. Sentana and Wadhwar(1992) found through empirical analysis that when volatility was low in the British and American markets, short-term stock returns showed a positive series correlation. However, when volatility is high, short-term stock returns will show negative serial correlation. And as volatility increases, positive feedback traders will have more influence on stock prices. Wang and Zhou(2009) found similar conclusions in the Chinese market through the empirical analysis of the stock market, that is, there is a reverse change relationship between the volatility of the stock market and the autocorrelation.

To sum up, positive feedback trading and its asymmetry are universal phenomena in the market. When the positive feedback transaction occurs, it will affect the volatility of stock prices, and the volatility will affect the correlation between stock returns, and ultimately affect the risk premium and expected return of stocks. However, when the retail investors chase up too fast, whether it will lead to too fast growth of the stock price bubble, resulting in subsequent stock price reversal decline; Or will rising volatility raise the risk premium and eventually lead to higher stock prices? At the same time, how will this ultimately affect expected returns in the event of a fall? The impact of positive feedback trade asymmetry on expected returns under the two conditions of rally and sell-off is still unclear, therefore, this paper will focus on the discussion of the above problems.

3 METHODOLOGY

From the Wind database, this paper selects the 20-year trading week data of 800 constituent stocks of China Securities Exchange from January 2003 to January 2023, totaling 800 stocks and 1,017 trading weeks. After excluding the data of late listed stocks with missing values, there are a total of 513838 effective weekly return rate data. In addition, weekly and annual data of the A-share index and the spot rate of 10-year Treasury bonds were used from March 2006 to December 2022.

3.1 Models

3.1.1 Model of Positive Feedback Transaction Asymmetry

According to Sentana and Wadhwar's (1992) model, there are two types of investors in the market: rational investors and positive feedback traders. The market shares of the two are respectively \( Q_t \) and \( Y_t \), which meet

\[
Q_t + Y_t = 1
\]  

(1)

Rational investors are always risk averse, so their share comes from the trade-off between risk and return:
Among, $E_{t-1}(r_t)$ is the investor's expected return of time $t$ at time $t-1$, $r_f$ is risk-free yield, $\sigma_t^2$ is Conditional fluctuation rate. Because the investors are risk averse, $\alpha_t > 0$ always established.

According to the positive feedback trading model built by Wan and Yang(2017), positive feedback traders trade according to the previous yield, buy when the stock rises, and sell when the stock falls. If the coefficient of pursuing the rise is $\gamma$, and the coefficient of selling down is $\gamma + \gamma_1$, then it is satisfied

$$Y_t = \gamma r_{t-1} + \gamma_1 r_{t-1} I[r_{t-1} > 0]$$ \hspace{1cm} (3)

$r_{t-1}$ is the previous yield. $I[r_{t-1} > 0]$ is virtual variables that distinguish between up and down, Taking 1 when the price rises and 0 when it falls. Therefore, $-\gamma_1$ can indicate the asymmetry of positive feedback trading. The greater the value, the greater the intensity of chasing up is stronger than killing the fall, and the same increase will lead to more trading volume.

By substituting equations (2) and (3) into Equation (1), the regression equation can be obtained:

$$r_t = \beta_0 + \beta_1 r_{t-1} + \beta_2 r_{t-1}^2 + \beta_3 \sigma_t^2 + \beta_4 r_{t-1} \sigma_t^2 + \beta_5 r_{t-1} I[\sigma_t^2 > 0] + u_t$$ \hspace{1cm} (4)

Then GARCH(1,1) model was used to estimate the conditional volatility $\sigma_t^2$, and Equation (4) was taken as the mean value equation, then obtained:

$$\begin{align*}
\{ & u_t = \sigma_t \epsilon_t, \\
\sigma_t^2 = \theta_0 + \theta_1 u_{t-1}^2 + \theta_2 u_{t-1}^2 
\}$$ \hspace{1cm} (5)

Parameters in equations (4) and (5) are obtained by using maximum likelihood estimation method. Among them, the asymmetry index of positive feedback transaction can be obtained by dividing the two.

### 3.1.2 CAPM Model

$$E(r_{it}) = \alpha_t + r_{ft} + \beta_{im} [E(r_{mi}) - r_{ft}] + e_{it}$$ \hspace{1cm} (6)

Among, $r_{ft}$ is the risk free rate at moment $t$, $E(r_{it})$ is the return rate of stock $i$ at moment $t$, $E(r_{mi})$ is the return rate of market portfolio at moment $t$, $\beta_{im}$ presents the sensitivity of the return rate of stock $i$ to market risk.

The CAPM model can be rewritten into

$$E(r_{it}) - r_{ft} = \alpha_t + \beta_{im} [E(r_{mi}) - r_{ft}] + e_{it}$$ \hspace{1cm} (7)
If the test result significantly has $\alpha_i \neq 0$, then it means that the market has a risk premium besides the market risk.

### 3.1.3 A-CAPM Model

Based on Zhou Fang and Zhang Wei(2011)'s improved LACAPM model, this paper added positive feedback trading asymmetry factor to the CAPM model to form a two-factor model named A-Capm (Asymmetry-CAPM) model:

$$E(r_{it}) = \alpha_i + r_{ft} + \beta_{im}[E(r_{mt}) - r_{ft}] + \beta_{ia}[ASY_{it} - \beta_{am}[E(r_{mt}) - r_{ft}]] + \epsilon_{it} \quad (8)$$

Among them, $ASY_{it}$ is positive feedback trading asymmetry factor of stock $i$ at year $t$, $\beta_{im}$ and $\beta_{ia}$ respectively represent the sensitivity of stock $i$'s return rate to market risk and positive feedback trading asymmetry. $\beta_{am}$ represents the sensitivity of positive feedback trading asymmetry to market risk.

The advantage of such improvement is that the possible correlation between market risk premium and positive feedback trading asymmetry is taken into account, and the premium of the part related to positive feedback trading asymmetry and market risk is attributed to market risk premium to eliminate this correlation, so as to more clearly discover the impact of positive feedback trading asymmetry on stock returns.

The above A-CAPM model can be rewritten into

$$E(r_{it}) - r_{ft} = \alpha_i + \beta_{im}[E(r_{mt}) - r_{ft}] + \beta_{ia}[ASY_{it} - \beta_{am}[E(r_{mt}) - r_{ft}]] + \epsilon_{it} \quad (9)$$

If the test result $\alpha_i$ is not significantly different from 0, it means that the model can explain the stock return well.

### 4 EMPIRICAL ANALYSIS

#### 4.1 Measurement of Positive Feedback Trading Asymmetry

First of all, this paper makes a descriptive statistical analysis of the weekly rate of return data of 800 constituent stocks from 2003 to 2022, the results are shown in Table 1.

**Table 1**: Descriptive statistics of 20 years' weekly return rate of CSI 800 constituent stocks.

<table>
<thead>
<tr>
<th>Stockcode</th>
<th>Mean</th>
<th>$\sigma$</th>
<th>min</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>000001.SZ</td>
<td>0.0034</td>
<td>0.0569</td>
<td>-0.2226</td>
<td>0.3576</td>
</tr>
<tr>
<td>000002.SZ</td>
<td>0.0053</td>
<td>0.0584</td>
<td>-0.2325</td>
<td>0.3310</td>
</tr>
<tr>
<td>000009.SZ</td>
<td>0.0047</td>
<td>0.0730</td>
<td>-0.2935</td>
<td>0.3867</td>
</tr>
<tr>
<td>00012.SZ</td>
<td>0.0039</td>
<td>0.0677</td>
<td>-0.2284</td>
<td>0.4928</td>
</tr>
<tr>
<td>000021.SZ</td>
<td>0.0033</td>
<td>0.0663</td>
<td>-0.3059</td>
<td>0.4983</td>
</tr>
<tr>
<td>000027.SZ</td>
<td>0.0028</td>
<td>0.0527</td>
<td>-0.3281</td>
<td>0.3182</td>
</tr>
<tr>
<td>000031.SZ</td>
<td>0.0033</td>
<td>0.0688</td>
<td>-0.3042</td>
<td>0.4667</td>
</tr>
</tbody>
</table>
According to the method of constructing the trading asymmetry index of positive feedback, the asymmetry index value of 800 constituent stocks in the past 20 years is obtained. As can be seen from Table 2, in the past 20 years, there has been a difference in the listing time of the current 800 constituent stocks. Over the 20-year period, eight of China Securities 800's positive feedback trading asymmetry indicators were positive and 12 were negative.

Table 2: Descriptive statistics of positive feedback trading asymmetry index.

<table>
<thead>
<tr>
<th>year</th>
<th>Mean</th>
<th>σ</th>
<th>min</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>-673429.1</td>
<td>1.22E+07</td>
<td>-2.05E+08</td>
<td>1.41E+07</td>
</tr>
<tr>
<td>2004</td>
<td>222.4684</td>
<td>3394.575</td>
<td>-5865.731</td>
<td>44503.07</td>
</tr>
<tr>
<td>2005</td>
<td>-469.7953</td>
<td>7247.681</td>
<td>-126814.1</td>
<td>11275.41</td>
</tr>
<tr>
<td>2006</td>
<td>-658634</td>
<td>1.19E+07</td>
<td>-2.14E+08</td>
<td>88301.04</td>
</tr>
<tr>
<td>2007</td>
<td>2094356</td>
<td>3.95E+07</td>
<td>-1165373</td>
<td>7.43E+08</td>
</tr>
<tr>
<td>2008</td>
<td>-31.40909</td>
<td>460.3631</td>
<td>-8421.11</td>
<td>1793.248</td>
</tr>
<tr>
<td>2009</td>
<td>20765.05</td>
<td>35405.93</td>
<td>-34732.09</td>
<td>7047710</td>
</tr>
<tr>
<td>2010</td>
<td>34.89687</td>
<td>12807.52</td>
<td>-157469.7</td>
<td>191130.1</td>
</tr>
<tr>
<td>2011</td>
<td>957.0581</td>
<td>43519.47</td>
<td>-399105.2</td>
<td>894570.7</td>
</tr>
<tr>
<td>2012</td>
<td>-13084.42</td>
<td>30097.82</td>
<td>-69997.52</td>
<td>37097.39</td>
</tr>
<tr>
<td>2013</td>
<td>1260299</td>
<td>2.94E+07</td>
<td>-3924.866</td>
<td>6.87E+08</td>
</tr>
<tr>
<td>2014</td>
<td>-114.1202</td>
<td>3171.3</td>
<td>-44974.58</td>
<td>39718.7</td>
</tr>
<tr>
<td>2015</td>
<td>-11.23797</td>
<td>858.0993</td>
<td>-11387.82</td>
<td>10525.56</td>
</tr>
<tr>
<td>2016</td>
<td>1.68E+07</td>
<td>4.16E+08</td>
<td>-24071.54</td>
<td>1.03E+10</td>
</tr>
<tr>
<td>2017</td>
<td>-122.2764</td>
<td>1211.874</td>
<td>-20195.14</td>
<td>7621.685</td>
</tr>
<tr>
<td>2018</td>
<td>-94.16075</td>
<td>2591.331</td>
<td>-38665.5</td>
<td>35482.82</td>
</tr>
<tr>
<td>2019</td>
<td>13786.72</td>
<td>10830.30</td>
<td>-1.41E+07</td>
<td>2.53E+07</td>
</tr>
<tr>
<td>2020</td>
<td>-3169.82</td>
<td>368897.3</td>
<td>-1.02E+07</td>
<td>50276.45</td>
</tr>
<tr>
<td>2021</td>
<td>-40.19519</td>
<td>321.5759</td>
<td>-6592.066</td>
<td>1759.1</td>
</tr>
<tr>
<td>2022</td>
<td>-52.5561</td>
<td>513.227</td>
<td>-11951</td>
<td>2026.443</td>
</tr>
</tbody>
</table>

As there are late listed stocks in the 800 constituent stocks of China Securities Exchange, the positive feedback trading asymmetry index is extremely large or small, and the existence of extreme value is often not of good practical significance. Therefore, in this paper, the value of the asymmetry factor is divided by its own largest value to show its distribution more intuitively. As can be seen from Figure 1, most of the positive feedback trading asymmetry...
factors are distributed on both sides of 0, indicating that the intensity of chasing and killing are generally comparable, and there may be significant differences in the short term.

Figure 1: Positive feedback trading asymmetric factor distribution.

This paper uses correlation analysis to verify the relationship between positive feedback trade asymmetry and retail trading intensity. If there is a significant positive correlation between positive feedback trading asymmetry and retail preference indicators, it can be considered that positive feedback trading asymmetry contains information related to retail trading intensity. Indicators selected in this paper to represent retail preference are: P/E ratio, volume, turnover rate and total market value.

According to the results of the correlation matrix shown in Table 3, the positive feedback trading asymmetry is positively correlated with the selected retail preference indicators, and the trading volume and turnover rate have a greater impact on the results of the positive feedback trading asymmetry, while the P/E ratio and total market value have a smaller impact on the results of the positive feedback trading asymmetry. This may be because positive feedback traders pay more attention to the historical data information and market sentiment that can be extracted from the stock market, and pay less attention to the operating conditions of the issuing companies themselves.

Table 3: Correlation between asymmetry and retail preference indicators.

<table>
<thead>
<tr>
<th></th>
<th>Asy</th>
<th>P/E</th>
<th>vol</th>
<th>turn</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asy</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P/E</td>
<td>0.0193</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vol</td>
<td>0.158*</td>
<td>-0.0944</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>turn</td>
<td>0.089*</td>
<td>0.0701</td>
<td>0.1062</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>0.0105</td>
<td>-0.0132</td>
<td>0.2257*</td>
<td>-0.1699*</td>
<td>1</td>
</tr>
</tbody>
</table>

Note: Table 3 reports the Pearson correlation coefficient between asymmetry and retail preference, and * indicates that it is significant at the 10% level.

4.2 Use CAPM to test risk premium in A-share market

This paper uses the spot rate of 10-year Treasury bonds as the risk-free rate of return and the A-share index as the market rate of return to test the CAPM model. The data of risk-free rate of return can be obtained from June 2006. The weekly rate of return data of China Securities 800 component stocks and the above variables were used to test the regression results of CAPM model.
As can be seen from the regression significance results in Table 4, most of the intercept term coefficients obtained by CAPM model regression are significantly different from 0, among which 688 are very significant, 34 are very significant and 30 are significant, while only 48 are not significantly different from 0. As can be seen from the regression results, the CAPM model can not explain the expected returns of the component stocks of China Securities 800 well, which means that there are other risks in the intercept term that cannot be explained by the market risk, that is, there are other risk premia besides the market risk.

### Table 4: Statistical significance results of CAPM model.

<table>
<thead>
<tr>
<th>Significance condition</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extremely significant(P&lt;0.001)</td>
<td>688</td>
</tr>
<tr>
<td>Very significant(P&gt;0.001&amp;P&lt;0.01)</td>
<td>34</td>
</tr>
<tr>
<td>Significant(P&gt;0.01&amp;P&lt;0.05)</td>
<td>30</td>
</tr>
<tr>
<td>Not significant(P&gt;0.05)</td>
<td>48</td>
</tr>
<tr>
<td>Total</td>
<td>800</td>
</tr>
</tbody>
</table>

### 4.3 Use A-CAPM to test risk premium in A-share market

In this paper, the improved CAPM model containing the positive feedback trading asymmetry factor is used to further regression test the return rate of China Securities 800. Since the value of the positive feedback trading asymmetry factor is often large, there is a magnitude difference between it and other variables in the model. In order to make the regression coefficient easier to observe and thus easier to summarize the rule, this paper carries out standardization processing on the positive feedback trade asymmetry factor:

\[
ASY_{tn} = \frac{Asy_{tn} - Asy_n}{\sigma_n}
\]

The positive feedback trading asymmetry factor after standardization is substituted into A-CAPM model for testing.

### Table 5: Statistical significance results of A-CAPM model.

<table>
<thead>
<tr>
<th>Significance condition</th>
<th>CAPM</th>
<th>A-CAPM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extremely significant(P&lt;0.001)</td>
<td>688</td>
<td>0</td>
</tr>
<tr>
<td>Very significant(P&gt;0.001&amp;P&lt;0.01)</td>
<td>34</td>
<td>0</td>
</tr>
<tr>
<td>Significant(P&gt;0.01&amp;P&lt;0.05)</td>
<td>30</td>
<td>13</td>
</tr>
<tr>
<td>Not significant(P&gt;0.05)</td>
<td>48</td>
<td>714</td>
</tr>
<tr>
<td>Total</td>
<td>800</td>
<td>727</td>
</tr>
</tbody>
</table>

In Table 5, due to the late listing time of some stocks, the number of asymmetric factors is small, and the regression test results lack value. Therefore, this paper finally obtained 727 effective regression results of A-CAPM. Among them, there are 714 intercept items that are not significant, and 13 intercept items that are significant. Compared with the CAPM model, A-CAPM model can well explain the expected returns of the constituent stocks of China Securities 800 after adding the positive feedback trading asymmetry factor into the model. It means that the positive feedback trading asymmetry factor can better explain the risk premium except market risk and has an impact on expected returns.
4.4 Testing the effect of positive feedback trading asymmetry factor on expected return

According to the above definition, when the positive feedback trade asymmetry factor is greater than 0, it means that the market's chasing strength is greater than the killing strength. When the value is less than 0, it means that the killing strength is greater than the chasing strength. In order to more directly test the impact of positive feedback trading asymmetry on expected earnings respectively under the two states of rally pursuit and sell-off, this paper divides the 800 constituent stocks of China Securities Securities into four categories according to the positive feedback trading asymmetry factor obtained: chasing>>killing, chasing>killing, chasing<killing, chasing<<killing. Each category contains 200 stocks.

Table 6 shows the regression coefficients of four categories of positive feedback trade asymmetry factors. It can be seen that when the chasing strength is much less than or less than the killing strength, the positive feedback trade asymmetry factor has a positive impact on the expected return, and the former has a greater impact than the latter. When the chasing strength is much greater than or greater than the killing strength, the positive feedback trading asymmetry factor has a negative impact on the expected return. Similarly, the former has a greater impact than the latter. From a practical point of view, when the intensity of chasing gains is greater than the intensity of killing losses, the stock market sentiment is high and the trading intensity of individual investors is increased. The continuous chasing gains trading is easy to lead to the bubble in the stock market, and the stock price is greater than its value, which eventually leads to the fall of the stock price. However, when the intensity of killing is greater than that of chasing, the sentiment of the stock market will be depressed, the trading intensity of institutions will increase, and the risk control strategy of institutions will require them to make selling measures, while the continuous decline of stock prices will lead to the breach of the risk control warning line among various institutions one by one, leading to the further decline of stock prices, until the situation that the stock price is lower than its value itself, eventually lead to the rise of stock prices.

Table 6: Descriptive statistics of positive feedback trading asymmetry categories.

<table>
<thead>
<tr>
<th>Category</th>
<th>Number</th>
<th>Mean</th>
<th>σ</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>C&lt;&lt;K</td>
<td>200</td>
<td>0.0341</td>
<td>0.2643</td>
<td>-0.5400</td>
<td>2.6552</td>
</tr>
<tr>
<td>C&lt;K</td>
<td>200</td>
<td>0.0243</td>
<td>0.2185</td>
<td>-1.0884</td>
<td>1.2617</td>
</tr>
<tr>
<td>C&gt;K</td>
<td>200</td>
<td>-0.0059</td>
<td>0.9690</td>
<td>-4.3759</td>
<td>11.720</td>
</tr>
<tr>
<td>C&gt;&gt;K</td>
<td>200</td>
<td>-0.0677</td>
<td>0.1969</td>
<td>-0.7365</td>
<td>0.5216</td>
</tr>
</tbody>
</table>

Where, C represents chasing strength, K represents killing strength.

5 CONCLUSIONS

This paper takes the weekly trading data of China Securities 800 constituent stocks as the research object, and constructs the positive feedback trading asymmetry index of each stock for 20 years. Pearson correlation test is used to analyze the correlation between the positive feedback trading asymmetry index and the retail preference index. The results show that the positive feedback trading asymmetry can represent the trading intensity of retail investors. By
using the methods of multiple regression and significance test, this paper tests the risk premium of the Chinese market, and confirms that the CAPM model cannot explain the expected rate of return of stocks well, and there is risk premium that cannot be explained by market risk in the market.

Therefore, this paper introduces the positive feedback trading asymmetry factor into the CAPM model to build the A-CAPM model. The significance results show that the improved model can better explain the expected return, and the positive feedback trading asymmetry factor has an important impact on the expected return of stocks. Finally, we classify the stocks according to the positive feedback trading asymmetry factor, and find that the positive feedback trading factor has different effects on the expected returns of stocks under the two states of chasing up and selling down. When the positive feedback trading factor is greater than 0 (chasing intensity is greater than killing intensity), the positive feedback trading factor has a negative impact on the expected return. When the positive feedback trading factor is less than 0 (the chasing intensity is less than the killing intensity), the positive feedback trading factor has a positive impact on the expected return, and the degree of impact on the expected return increases with the increase of the factor in both cases.

REFERENCES

Digital Transformation, Dynamic Capability and Green Technology Innovation: Empirical Evidence Based on Text Analysis Methods

Jiarun He¹, Binyi Sun¹,ᵇ

cnhjr1234@163.comᵃ, bysun@dbm.ecnu.edu.cnᵇ

East China Normal University, Shanghai, China¹

Abstract: Digital transformation is one of the important ways for enterprises to enhance their core competitiveness and achieve sustainable development. Based on the data of listed resource-based enterprises in China from 2012 to 2019, this paper uses python and crawler technology to collect the key words of "digital transformation" in the annual report of enterprises, and combines the dynamic capability theory and social network theory to empirically test the impact of digital transformation on enterprises' green technology innovation and its transmission mechanism. The research found that digital transformation can significantly improve the level of green technology innovation of enterprises; Digital transformation has a significant positive impact on innovation ability, growth ability and resource integration ability; Innovation ability, growth ability and resource integration ability all have a partial intermediary effect between digital transformation and green technology innovation. The research conclusion provides research ideas for the follow-up resource-based enterprises to promote green technology innovation and guide the digital transformation of enterprises.

Keywords: Digital Transformation, Text Recognition, Dynamic Capability, Green Technology Innovation.

1 INTRODUCTION

With the development of big data, cloud computing and other technologies, the digital economy has gradually become an important support for high-quality economic development. The progress of digital economy can reduce the consumption of social resources and energy, and consider economic and ecological benefits. Under this background, resource-based enterprises with the characteristics of serious environmental damage, low resource utilization and high energy consumption complete the digital transformation and promote green technology innovation has profound significance.

The application of digital technology can reshape the internal management mechanism of enterprises¹⁸, promote the intelligent production management of enterprises, and enhance the collaborative innovation ability of enterprises⁹. The progress of digital technology has created opportunities for enterprise innovation, but only a few scholars have linked digitization with enterprise green technology innovation. According to the literature review, the existing research on the mechanism of digital transformation and green technology innovation of enterprises is
not sufficient, and at the same time, the research on dynamic capability of enterprises is also ignored. The dynamic capability is the ability to organize and coordinate internal and external resources to respond more flexibly to the unstable external environment [21]. Based on this, it is of profound practical significance to study the relationship between digital transformation and green technology innovation of listed resource-based enterprises in China.

The possible research contributions of this paper are as follows: on the one hand, it explores the mechanism between digital transformation and green technology innovation, and reveals the internal mechanism of dynamic capability of enterprises; On the other hand, it provides constructive solutions for green technology innovation of resource-based enterprises and provides reference for subsequent relevant research.

2 LITERATURE REVIEW

2.1 Dynamic Capability

Teece et al. (1997) first believed that dynamic capability is the ability of enterprises to integrate, build and reconfigure internal and external resources. Dynamic capability includes adaptability, absorption, and innovation. Adaptability refers to the ability of an organization to identify and seize opportunities, and absorptive ability refers to the ability of an organization to identify, absorb and apply new information, and innovative ability represents the ability of an organization to create new products and explore new markets [24]. Based on the actual situation of Chinese enterprises, scholars have different views. For example, He et al. (2006) believed that dynamic capability can be divided into market potential, organizational flexibility, strategic isolation, organizational learning, and organizational change; Luo et al. (2009) found that dynamic ability consists of absorption ability, perception ability, relationship ability and integration ability. Dynamic capability is a high-level capability of an organization, which can change ordinary capabilities [28]. The simple application of digital technology is a shallow capability that is easy to be imitated [16], but digital transformation can promote the improvement of dynamic capability, enabling enterprises to quickly adjust their production and operation activities to meet the challenges brought by the changes in the internal and external environment [13].

2.2 Digital Transformation and Green Technology Innovation

According to the social network theory, social networks breed innovative activities [17]. When organizations are faced with development bottlenecks, they usually choose to seek cooperation with organizations that complement their own capabilities through social networks (Croft et al. 2006). In the era of digital economy, digitalization has changed the technological and economic environment and the operation of social institutions [12], and social networks have expanded rapidly.

Green technology innovation is a series of creative activities that follow ecological laws, reduce environmental pollution and energy and resource consumption, and achieve sustainable development [11]. On the one hand, resource-based enterprises face the problem of high technical complexity and weak independent innovation capability. On the other hand, they face high risks due to the impact of policy and market uncertainty [6]. In this context, digitalization can help
resource-based enterprises greatly expand social networks, share technology dividends \(^4\), and spread risks through cooperation. The following assumption is derived:

Hypothesis 1. Digital transformation is positively affecting green technology innovation.

2.3 The Intermediary Role of Dynamic Capability

Digital technology is radically changing business networks and key inter-enterprise relationships \(^5\). Digital transformation is not only the process of updating the production mode of the organization, but also the process of building dynamic capability for the continuous strategic update of the organization \(^26\). Using the digital infrastructure such as the Internet of Things and the platform based on artificial intelligence, enterprises can collect and analyse large-scale real-time data for forecasting \(^7\), greatly improving the dynamic ability of enterprises to respond to environmental changes. The following assumption is derived:

Hypothesis 2. Digital transformation is positively affecting dynamic capability.

According to the theory of dynamic capability, dynamic capability is the core driving factor of strategic renewal \(^3\), which can improve the green innovation performance of enterprises \(^15\). Therefore, enhancing dynamic capability is the strategic choice for the transformation and upgrading of resource-based enterprises in the future. With reference to Teece et al. (1997), Wang et al. (2007) and other studies, this paper divides dynamic capability into innovation capabilities, growth capabilities and resource integration capabilities. Among them, the integration ability is the ability to achieve sustainable and green development through operation and integration of resources \(^20\); Innovation capability is the ability of enterprises to explore innovative technologies and unknown fields \(^30\) and respond quickly to changes in cutting-edge information of green innovative technologies; The growth ability lays the material foundation for the long-term green technological innovation of enterprises. Lack of resources will have a negative impact on innovation \(^1\). Digital transformation improves the ability of integrating resource. Stakeholders of green technology innovation can achieve collaborative research and development and resource connectivity by building innovation alliances. Based on this, the following assumptions are made:

Hypothesis 3a. Innovation ability positively affects green technology innovation;

Hypothesis 3b. The ability of resource integration has a positive impact on green technology innovation;

Hypothesis 3c. Growth ability is positively influencing green technology innovation.

As mentioned above, digital transformation has improved the dynamic capability of resource-based enterprises. On the one hand, the ability of enterprises to utilize resources is enhanced, and on the other hand, it provides new opportunities for enterprises to explore green technology innovation. It can be seen that digital transformation fosters dynamic capability of enterprises to promote green technology innovation, and dynamic capability plays an intermediary role between the two. To sum up, the following assumptions are made:

Hypothesis 4. Digital transformation has a positive impact on green technology innovation through dynamic capability.
Based on the above discussion, the theoretical research framework of this paper is constructed, as shown in Figure 1:

![Figure 1: The theoretical framework of this paper.](image)

### 3 METHODOLOGY

#### 3.1 Data Collection

Resource-based industries can be divided into narrow sense and broad sense. With reference to the relevant literature \[^{19, 25}\], this paper chooses its narrow definition, namely, the industry of mining, washing and primary processing of energy and mineral resources.

Considering the availability of data and the convenience of research, this paper takes the resource-based enterprises listed in Shanghai and Shenzhen A-shares in China from 2012 to 2019 as a sample, and excludes the enterprises with ST and *ST and those with serious data loss. Enterprise financial indicator data comes from CSMAR database, and green patent data comes from CNRDS database. After screening, 250 enterprises with a total of 2000 observations were finally determined. In addition, in order to reduce the impact of extreme values on the model, the upper and lower 1% quantiles of all continuous variables are shrunk.

#### 3.2 Measures

##### 3.2.1 Interpreted Variable

Green technology innovation level. Green invention patents can measure the level of green technological innovation of enterprises. With reference to the research of Tang et al. (2022), this paper selects the number of green patent applications add 1 and then takes the logarithm as the measurement index of green technological innovation of enterprises.

##### 3.2.2 Explanatory Variable

Digital transformation level. With the progress of natural language processing (NLP) and other technologies, text analysis has been used by scholars in the fields of economics and management in the past two years. Referring to Wu Fei et al. (2021) (Wu et al. Further, the word frequency of the words related to "digital technology application" in the annual report of the sample enterprises is counted. In order to improve the accuracy and matching of word frequency analysis, this paper constructs a digital word segmentation dictionary from six dimensions: intelligent production, mobile internet, digital technology, data management, modern
information system, and universal digitalization, and then carries out word frequency statistics. Finally, the frequency of the calculated digital technology application-level word is processed with logarithm to make its distribution more stable.

3.2.3 Intermediary Variable

As mentioned above, this paper selects innovation ability, resource integration ability and growth ability as the dynamic capacity proxy variables of this study. Among them, the innovation ability is measured by the ratio of R&D expenditure to operating income with reference to the research of Guo et al., (2019); The resource integration capacity is measured by the ratio of net profit to the total assets of the enterprise; For growth ability, refer to Zhou (2014) and other studies, and use sustainable growth rate to measure the sustainable development ability of resource-based enterprises. Sustainable growth rate = return on net assets * earnings retention rate / (1 - return on net assets * earnings retention rate) = (net profit / total closing balance of owner's equity) * [1 - pre-dividend tax per share / (current value of net profit / current value of paid-in capital)] / (1 - numerator).

3.2.4 Control Variable

Referring to relevant research, this paper selects control variables including enterprise size, enterprise age, etc. The variables selected in this paper are described in Table 1 below:

<table>
<thead>
<tr>
<th>Variable Type</th>
<th>Variable Name</th>
<th>Symbolic</th>
<th>Variable Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control variable</td>
<td>Equity concentration</td>
<td>CR10</td>
<td>Sum of shareholding ratio of top ten shareholders</td>
</tr>
<tr>
<td></td>
<td>Enterprise age</td>
<td>Age</td>
<td>The natural logarithm of the current year minus the year of listing</td>
</tr>
<tr>
<td></td>
<td>Leverage</td>
<td>Leverage</td>
<td>Total liabilities/assets</td>
</tr>
<tr>
<td></td>
<td>Enterprise nature</td>
<td>Soe</td>
<td>Soe=1, Others= 0</td>
</tr>
<tr>
<td></td>
<td>Enterprise size</td>
<td>Size</td>
<td>The natural logarithm of the total assets of the enterprise</td>
</tr>
<tr>
<td></td>
<td>Proportion of fixed assets</td>
<td>Tang</td>
<td>Net fixed assets/total assets</td>
</tr>
</tbody>
</table>

3.3 Model Settings

In order to test the research hypothesis proposed in this paper, the following three models are finally set:

\[ G_{\text{Innovation}}_{it} = \beta_0 + \beta_1DI + \Sigma \beta_mControl_{it} + \Sigma \text{Year} + \Sigma \text{Industry} + \epsilon_{i,t} \]  
\[ Med_{it} = \beta_0 + \beta_1DI + \Sigma \beta_mControl_{it} + \Sigma \text{Year} + \Sigma \text{Industry} + \epsilon_{i,t} \]  
\[ G_{\text{Innovation}}_{it} = \beta_0 + \beta_1DI + \beta_2Med_{it} + \Sigma \beta_mControl_{it} + \Sigma \text{Year} + \Sigma \text{Industry} + \epsilon_{i,t} \]
Among them, \( i \) and \( t \) represent enterprises and time respectively, and GInnovation is the explained variable, representing the level of green technology innovation of enterprises. DI is the core explanatory variable, representing digital transformation level. Med includes three intermediary variables, namely innovation ability, growth ability and resource integration ability. Control indicates a series of control variables, such as enterprise scale. \( \varepsilon \) Represents the random error term of the model.

4 RESULTS AND DISCUSSION

4.1 Descriptive Statistical Analysis

Table 2 below shows the descriptive statistics of the main variables. The maximum value of GInnovation is 4.2973, the minimum value is 0, and the standard deviation is 1.0469. There is a large difference in the level of green technology innovation among enterprises, which is basically consistent with the existing research results. The maximum value of DI is 4.4128, the minimum value is 0, and the standard deviation is 1.0112, which also reflects the large differences in the application of digital technology among enterprises, and is consistent with the slow digital transformation of resource-based enterprises. The maximum value of innovation capacity is 0.0795, the average value is 0.0225, and the median value is 0.0209, which indicates that a considerable number of resource-based enterprises have innovation levels above the average level, reflecting the overall high innovation capacity.

<table>
<thead>
<tr>
<th>Variables</th>
<th>N</th>
<th>Mean</th>
<th>S.D.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>GInnovation</td>
<td>2000</td>
<td>0.829</td>
<td>1.047</td>
<td>0.000</td>
<td>4.297</td>
</tr>
<tr>
<td>DI</td>
<td>2000</td>
<td>1.893</td>
<td>1.011</td>
<td>0.000</td>
<td>4.413</td>
</tr>
<tr>
<td>Age</td>
<td>2000</td>
<td>2.403</td>
<td>0.590</td>
<td>0.693</td>
<td>3.239</td>
</tr>
<tr>
<td>CR10</td>
<td>2000</td>
<td>58.906</td>
<td>15.286</td>
<td>23.675</td>
<td>92.515</td>
</tr>
<tr>
<td>Tang</td>
<td>2000</td>
<td>0.342</td>
<td>0.172</td>
<td>0.036</td>
<td>0.788</td>
</tr>
<tr>
<td>Size</td>
<td>2000</td>
<td>22.831</td>
<td>1.405</td>
<td>20.271</td>
<td>26.602</td>
</tr>
<tr>
<td>Soc</td>
<td>2000</td>
<td>0.535</td>
<td>0.499</td>
<td>0.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Leverage</td>
<td>2000</td>
<td>0.476</td>
<td>0.192</td>
<td>0.093</td>
<td>0.908</td>
</tr>
<tr>
<td>Innovation</td>
<td>2000</td>
<td>0.023</td>
<td>0.018</td>
<td>0.000</td>
<td>0.080</td>
</tr>
<tr>
<td>Growth</td>
<td>2000</td>
<td>0.035</td>
<td>0.101</td>
<td>-0.411</td>
<td>0.397</td>
</tr>
<tr>
<td>Integration</td>
<td>2000</td>
<td>0.029</td>
<td>0.051</td>
<td>-0.164</td>
<td>0.184</td>
</tr>
</tbody>
</table>

The correlation analysis shows that DI, Innovation, Growth, and Integration are all related to GInnovation, and through the significance test at the level of 1%, the variable selection is representative; At the same time, the level of digital transformation is also significantly related to innovation ability, growth ability and resource integration ability, and has passed the significance test of 1% level. Further, the variance expansion factor of all variables is calculated, and the mean value is 2.14, which is far less than 10, indicating that there is no multicollinearity problem.
4.2 Regression results

The results of regression analysis are shown in Table 3 and Table 4 below. Model (1) is the benchmark regression model of this paper, which only adds all control variables. Model (2) adds digital transformation level (DI) to the benchmark regression model, and it can be found that the coefficient is 0.096, and it has passed the significance test at the level of 1%. Hypothesis 1 has been verified, and the improvement of digital transformation level is conducive to improving the level of green technology innovation of enterprises.

Table 3: Regression analysis results (Model 1-4)

<table>
<thead>
<tr>
<th>Explained variable</th>
<th>GLnno</th>
<th>GLnno</th>
<th>Inno</th>
<th>GLnno</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DI</td>
<td>0.096***</td>
<td>0.003***</td>
<td>0.083***</td>
<td></td>
</tr>
<tr>
<td>Innovation</td>
<td></td>
<td></td>
<td></td>
<td>4.997***</td>
</tr>
<tr>
<td>Growth</td>
<td>-0.078*</td>
<td>-0.114**</td>
<td>-0.005***</td>
<td>-0.089*</td>
</tr>
<tr>
<td>Age</td>
<td>(-1.73)</td>
<td>(-2.50)</td>
<td>(-6.16)</td>
<td>(-1.94)</td>
</tr>
<tr>
<td>CR10</td>
<td>-0.002</td>
<td>-0.002</td>
<td>-0.000***</td>
<td>-0.001</td>
</tr>
<tr>
<td>Tang</td>
<td>(-1.23)</td>
<td>(-1.20)</td>
<td>(-2.94)</td>
<td>(-0.94)</td>
</tr>
<tr>
<td>Size</td>
<td>-0.240*</td>
<td>-0.195</td>
<td>0.001</td>
<td>-0.202</td>
</tr>
<tr>
<td>Soe</td>
<td>(-1.81)</td>
<td>(-1.47)</td>
<td>-0.6</td>
<td>(-1.53)</td>
</tr>
<tr>
<td>Leve</td>
<td>0.420***</td>
<td>0.400***</td>
<td>-0.003***</td>
<td>0.414***</td>
</tr>
<tr>
<td>Constant</td>
<td>-21.25</td>
<td>-19.87</td>
<td>(-8.11)</td>
<td>-20.34</td>
</tr>
<tr>
<td>Observations</td>
<td>2,000</td>
<td>2,000</td>
<td>2,000</td>
<td>2,000</td>
</tr>
<tr>
<td>R²</td>
<td>0.258</td>
<td>0.266</td>
<td>0.213</td>
<td>0.272</td>
</tr>
</tbody>
</table>
### Table 4: Regression analysis results (Model 5-8)

<table>
<thead>
<tr>
<th>Explained variable</th>
<th>Growth</th>
<th>Glnno</th>
<th>Integ</th>
<th>Glnno</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Explanatory variable</strong></td>
<td>Model (5)</td>
<td>Model (6)</td>
<td>Model (7)</td>
<td>Model (8)</td>
</tr>
<tr>
<td>DI</td>
<td>0.004*</td>
<td>0.095***</td>
<td>0.002*</td>
<td>0.094***</td>
</tr>
<tr>
<td>Innovation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Growth</td>
<td>0.474**</td>
<td></td>
<td></td>
<td>1.083**</td>
</tr>
<tr>
<td>Age</td>
<td>0.011**</td>
<td>-0.119***</td>
<td>0.005**</td>
<td>-0.119***</td>
</tr>
<tr>
<td>CR10</td>
<td>0</td>
<td>-0.002</td>
<td>0</td>
<td>-0.002</td>
</tr>
<tr>
<td>Tang</td>
<td>-0.026*</td>
<td>-0.182</td>
<td>-0.012*</td>
<td>-0.182</td>
</tr>
<tr>
<td>Size</td>
<td>0.017***</td>
<td>0.392***</td>
<td>0.009***</td>
<td>0.390***</td>
</tr>
<tr>
<td>Soe</td>
<td>-0.020***</td>
<td>0.185***</td>
<td>-0.016***</td>
<td>0.193***</td>
</tr>
<tr>
<td>Leve</td>
<td>-0.157***</td>
<td>-0.471***</td>
<td>-0.122***</td>
<td>-0.412***</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.286***</td>
<td>-7.728***</td>
<td>-0.138***</td>
<td>-7.714***</td>
</tr>
<tr>
<td>Observations</td>
<td>2,000</td>
<td>2,000</td>
<td>2,000</td>
<td>2,000</td>
</tr>
<tr>
<td>R²</td>
<td>0.09</td>
<td>0.268</td>
<td>0.208</td>
<td>0.268</td>
</tr>
</tbody>
</table>

Further, this paper uses the research of Wen et al. (2014) for reference to test the intermediary effect. Model (3) tests the relationship between digital transformation level and innovation capability, and the result shows that there is a significant positive correlation (1% level) between the two. Model (4) adds digital transformation level and innovation capability to the benchmark model at the same time. It can be seen from the table that the regression coefficient of digital transformation decreases from 0.096 to 0.083, but it is still significant, indicating that there is a mediating effect between innovation capability and digital transformation level and green technology innovation. Hypothesis 2 Assumption 3a holds; The model (5) test found that there was a significant correlation between growth ability and digital transformation. The model (6) that included both digital transformation level and growth ability in the benchmark regression model found that the regression coefficient was still significant (at the level of 1%) before digital
transformation, and the regression coefficient was reduced, so the hypothesis 3c was established; The model (7) test found that there is a significant correlation between the resource integration capability and the digital transformation level. After the digital transformation and resource integration capability are added, the regression coefficient before the digital transformation level is still significant (1% level). Therefore, assuming that 3b is established, the resource integration capability has a mediating effect between the digital transformation and green technology innovation.

In order to enhance the reliability of the intermediary effect test, Sobel test and nonparametric percentile Bootstrap test are conducted in this paper. Sobel test results show that Z value is 3.45, which is significant at 1% level. The Bootstrap sampling method is set to 5000 times. The results show that the 95% level confidence interval of the direct effect is (0.006, 0.0202), and the 95% level confidence interval of the indirect effect is (0.041, 0.125). It can be seen that the confidence interval of both effects does not contain 0, and both effects pass the 1% level significance test, which once again proves that the intermediary effect exists.

4.3 Robustness

In order to further increase the explanatory power of the model, this paper makes the following robustness tests: (1) The explained variable lags. The green patent itself has a certain time lag effect, so the green technology innovation level is added to the model after a period of lag, and the results are basically consistent with the previous research conclusions; (2) Tendency score matching method (PSM). Influenced by many factors, the level of green technology innovation, the level of digital transformation, and the dynamic capacity variables may have a causal relationship with each other, and the existence of self-selection of samples may lead to the existence of endogenous problems in the model. This paper uses the method to further test. The results showed that the t-test before unmatching was significant (7.98), and the average processing effect ( ) after matching was still significant (t statistic value was 3.00).

5 CONCLUSIONS

Based on the sample of resource-based enterprises listed in Shanghai and Shenzhen A shares in China, this paper empirically tests that digital transformation has a positive effect on improving the level of green technology innovation of enterprises. It uses the sequential test method, Sobel test and Bootstrap method to test the dynamic capability including innovation ability, growth ability and resource integration ability to play an intermediary role between digital transformation and green technology innovation. It reveals the internal mechanism of the impact of digital technology on green technology innovation. Finally, the following conclusions are drawn: the digital transformation has significantly improved the level of green technology innovation of enterprises. The specific transmission mechanism is that enterprises can improve the application level of digital technology to enhance the higher level of dynamic capability, enable technological innovation, and improve the level of green technology innovation.

Resource-based enterprises should accelerate the digital transformation, introduce green and intelligent production and manufacturing equipment, and comprehensively promote the deep integration of digital technology and green technology innovation, to improve the efficiency of resource and energy utilization and achieve win-win economic and ecological benefits. In
addition, make full use of digital technology to break information barriers and improve dynamic capability.

REFERENCES


Big Data Evaluation on Blue Carbon Economic Development Level in Daguang Bay, Jiangmen City

Jianhua Xiao¹, Wanqing Zhu¹*

Corresponding author and e-mail: Wanqing Zhu, 3624385176@qq.com

School of Economics and Management, Wuyi University, Jiangmen, China¹

Abstract: In order to promote the development of blue carbon economy in Daguang Bay of Jiangmen City, realize the blue rise and promote the high-quality development of Marine economy, this paper constructs the evaluation index system of blue carbon economy development in Daguang Bay based on the relevant Marine data of Jiangmen City from 2010 to 2019. The entropy method is adopted to calculate the weight of the index and analyze the development situation of blue carbon economy in Daguang Bay of Jiangmen City from the perspective of time. The results show that the development level of blue carbon economy in Daguang Bay of Jiangmen City, including Xinhui District and Enping City, is on the rise, while that in Taishan City is on the decline. However, the development level of blue carbon economy in Xinhui District and Taishan City is better than that in Enping City. It is found that in order to achieve high-quality development of Marine economy in Jiangmen, there are still problems such as poor quality of Marine environment, insufficient driving force of Marine innovation and unbalanced regional development. Therefore, it is proposed to strengthen Marine environmental protection, promote Marine science and technology innovation, and focus on increasing the development of blue carbon economy in Taishan City and Enping City.

Keywords: Daguang Bay, Blue Carbon Economy, Entropy Method.

1 INTRODUCTION

The rich resources contained in the ocean are an important support for social and economic development. The United Nations once pointed out in 2001 that "the 21st century is the century of the ocean", and countries all over the world have listed the development of the Marine economy, the protection of the Marine environment and the protection of their own maritime rights and interests as important development strategies [1]. As a maritime power, the report of the 19th National Congress of the Communist Party of China clearly requires that we adhere to land and sea coordination and accelerate the construction of a maritime power, which once again sounded the clarion call for building a maritime power. At the same time, as a major carbon dioxide emitter, China will actively participate in global environmental governance and fulfill its emission reduction commitments. "Strive to reach the peak of carbon dioxide emissions by 2030, and strive to achieve carbon neutrality by 2060" is a major strategic decision made by the Party Central Committee and an important task for governments at all levels to implement new development concepts and promote high-quality development.
As early as 2015, the Central Committee of the Communist Party of China and the State Council issued the "Overall Plan for the Reform of the Ecological Civilization System" and pointed out that "it is necessary to establish an effective mechanism for ocean carbon sink and accelerate the expansion of the blue economic space." Blue carbon sink (abbreviated as "blue carbon"), that is, ocean carbon sink, is the process, activity and mechanism of using marine organisms to absorb carbon dioxide in the atmosphere and fix it in the ocean [6]. As a new idea of low-carbon economic development, blue carbon sink has been studied by many scholars. Hu Jianbo and Zhang Qiang [3] defined the connotation of blue carbon sink, and put forward corresponding countermeasures and suggestions for promoting the development of Blue carbon sink in China to better promote the realization of emission reduction targets based on its development status and problems. Wang Chengrong [9] used the 21st Century Maritime Silk Road as the background to analyze the importance of the development of blue carbon in Guangdong Province to control greenhouse gas emissions, marine ecological environmental degradation, and the implementation of the blue carbon "Belt and Road", and explore its development path. Macreadie PI et al [4] summarized 10 important issues about blue carbon to improve blue carbon science and its related applications in mitigating climate change. Moritsch Monica M et al. [5] explored the maximum benefits of blue carbon sequestration under four management scenarios: management retreat, management retreat plus dam removal, erosion in high-risk areas, and erosion in moderate-to-high-risk areas, and compared their economic values.

As a major marine city in Guangdong Province, Jiangmen City closely follows the national and provincial strategic deployment, vigorously promotes the construction of a "strong marine economy", fully promotes the comprehensive development of the marine economy, and plans a "blue rise". For Jiangmen to realize the blue rise and promote the high-quality development of the marine economy, the scientific development of Daguang Bay is the only way [8]. The development of blue carbon economy can effectively promote the marine ecological balance of Daguang Bay and help to promote the establishment of carbon financial market, and it is also of great significance to promote the development of new marine business in Daguang Bay. Therefore, the development of marine carbon sinks is the material basis for Daguang Bay to implement the concept of green development, build a marine ecological civilization, and achieve high-quality development of the marine economy.

2 CONSTRUCTION AND RESEARCH METHODS OF BIG DATA EVALUATION INDEX SYSTEM FOR BLUE CARBON ECONOMIC DEVELOPMENT IN DAGUANG BAY, JIANGMEN CITY

2.1 Construction of Big Data Evaluation Index System for Blue Carbon Economic Development in Daguang Bay, Jiangmen City

In 2013, the Guangdong Provincial Government approved the scope of the Daguang Bay Comprehensive Development Economic Zone [2]. Based on the availability of data, this article takes Xinhui District, Taishan City and Enping City as the research objects to explore the blueprint of the three.
In order to objectively and scientifically evaluate the development level of the blue carbon economy in Daguang Bay of Jiangmen City, it is necessary to establish a set of evaluation index system, and to integrate previous researches and establish an index system on the basis of literature analysis, as shown in Table 1. The original indicator data comes from the "Jiangmen Statistical Yearbook" (2010-2019).

Table 1: Construction of Big Data evaluation index system for blue carbon economic development in Daguang Bay, Jiangmen City

<table>
<thead>
<tr>
<th>Target layer</th>
<th>Criterion layer</th>
<th>Element layer</th>
<th>Dimension</th>
<th>Indicator attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>The development level of the blue carbon economy of Daguang Bay</td>
<td>Carbon sink indicators (A1)</td>
<td>Main energy consumption (B1)</td>
<td>Tons of standard media</td>
<td>Inverse</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Forest cover rate (B2)</td>
<td>%</td>
<td>Positive</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Green area rate of built-up area (B3)</td>
<td>%</td>
<td>Positive</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Green coverage area (B4)</td>
<td>Hectares</td>
<td>Positive</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Forest stock (B5)</td>
<td>Ten thousand cubic meters</td>
<td>Positive</td>
</tr>
<tr>
<td></td>
<td>Marine economic development indicators (A2)</td>
<td>GDP per capita (B6)</td>
<td>Yuan</td>
<td>Positive</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Added value of tertiary industry (B7)</td>
<td>Ten thousand yuan</td>
<td>Positive</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Freshwater aquaculture area (B8)</td>
<td>Mu</td>
<td>Positive</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Aquatic product output (B9)</td>
<td>Ton</td>
<td>Positive</td>
</tr>
</tbody>
</table>

2.2 Research Methods

Entropy is a measure of uncertainty. Entropy method is a mathematical method used to judge the degree of dispersion of a certain index. The greater the degree of dispersion, the greater the impact of this indicator on the evaluation. The calculation process of the entropy method evaluation model is as follows:

- Construct the original index data matrix. Suppose the original index evaluation matrix for the high-quality development of marine economy in Guangdong Province is:

\[
X_{ij} = \begin{bmatrix}
    x_{11} & x_{12} & \cdots & x_{1j} \\
    x_{21} & x_{22} & \cdots & x_{2j} \\
    \vdots  & \vdots  & \ddots & \vdots  \\
    x_{i1} & x_{i2} & \cdots & x_{ij}
\end{bmatrix}
\]  

(1)

Formula (1): represents the original evaluation index matrix, and represents the index data of the th index in the th year.
Dimensionless processing of data. Since most of the measurement units and orders of magnitude of $X_{ij}$ are different, it is necessary to standardize each indicator to eliminate the impact of different dimensions on the results.

\[
X'_{ij} = \frac{x_{ij} - \min[x_{ij}]}{\max[x_{ij}] - \min[x_{ij}]}
\]  

(2)

\[
X'_{ij} = \frac{\max[x_{ij}] - x_{ij}}{\max[x_{ij}] - \min[x_{ij}]}
\]  

(3)

Among them, the positive index adopts formula (2), and the negative index adopts formula (3). $X_{ij}$ is a standardized value, and $\min[x_{ij}]$ and $\max[x_{ij}]$ are the minimum and maximum values of a certain index respectively. Formula (4) is the standardized evaluation matrix:

\[
A = \begin{bmatrix}
a_{11} & a_{12} & \cdots & a_{1j} \\
a_{21} & a_{22} & \cdots & a_{2j} \\
\vdots & \vdots & \ddots & \vdots \\
a_{i1} & a_{i2} & \cdots & a_{ij}
\end{bmatrix}
\]  

(4)

$A$ is the standardized evaluation matrix, and $a_{ij}$ is the standardized index value.

- Calculate the contribution degree $Y_{ij}$ of the evaluation index. After the data is dimensionlessly processed, the contribution of the $j$-th indicator in the $i$-th year is calculated. The formula is shown in (5):

\[
Y_{ij} = \frac{X'_{ij}}{\sum_{i=1}^{m} X'_{ij}}
\]  

(5)

- Calculate the index information entropy $e_j$. Used to reflect the amount of information of the $j$-th index. $k = 1/\ln(m)$, $m$ is the evaluation year. In the evaluation system for the high-quality development of marine economy in Guangdong Province constructed in this paper, $m$ is 9. The calculation formula is shown in (6):

\[
e_j = -k \sum_{i=1}^{m} (Y_{ij} \times \ln Y_{ij})
\]  

(6)

- Calculate the information utility value. Its value directly affects the size of the weight. The greater the information utility value, the greater the importance of the evaluation and the greater the weight. As shown in formula (7):

\[
d_j = 1 - e_j
\]  

(7)
Determine the weight of the evaluation index. The greater the weight, the greater the contribution to the evaluation result. As shown in formula (8):

\[ W_j = \frac{d_j}{\sum_{j=1}^{m} d_j} \]  

(8)

Calculate the comprehensive evaluation value of the sample. Use the product of the j-th index weight \( W_j \) and the standardized matrix \( A \) as the evaluation value \( F_{ij} \), that is, \( F_{ij} = W_j \times A \).

3 BIG DATA EVALUATION OF BLUE CARBON ECONOMIC DEVELOPMENT IN DAGUANG BAY, JIANGMEN CITY

3.1 Analysis on the Weights of Blue Carbon Economic Development Indexes in Daguang Bay, Jiangmen City

According to the results in Table 2, on the whole, carbon sink sink has a greater impact on blue carbon economic development (A1) than Marine economic development (A2) in Daguang Bay, Jiangmen City. Specifically, in 2010, the gre-en coverage area (B4) and forest stock volume (B5) have a higher weight (0.174 and 0.152, respectively), indicating that the green coverage area has a significant impact on carbon during the development of carbon sink in Daguang Bay, Jiangmen City. The achievement of the emission reduction target has the greatest impact, followed by the forest stock volume, and the forest coverage rate (B2) has the lowest weight of 0.076. For indicators related to marine economic development, the added value of the tertiary industry (B7) has the highest weight and has the most profound impact, followed by per capita GDP (B6), and freshwater aquaculture area (B8) has the lowest weight. In 2019, the biggest impact on the development of Daguang Bay’s carbon sinks is still the green coverage area (B4), with a weight of 0.201, and the biggest impact on the development of Daguang Bay’s marine economy is still the added value of the tertiary industry (B7).

On the whole, the development of the blue carbon economy of Daguang Bay in Jiangmen City has the same presentation at different development stages. The green coverage area and the added value of the tertiary industry are the main fac-tors affecting the development of the blue carbon economy in Daguang Bay, Jiangmen. Highlight the important status of both. In the future, while developing the marine economy, we can increase the area of green coverage, optimize the structure of the tertiary industry, expand the added value, and promote the high-quality development of the marine economy of Daguang Bay.

Table 2:2010-2019 Daguang Bay Blue Carbon Economic Development Index Weight in Jiangmen City

<table>
<thead>
<tr>
<th>index</th>
<th>year</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
<th>2014</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B1</td>
<td>0.078</td>
<td>0.076</td>
<td>0.075</td>
<td>0.072</td>
<td>0.074</td>
<td></td>
</tr>
<tr>
<td>B2</td>
<td>0.076</td>
<td>0.075</td>
<td>0.076</td>
<td>0.071</td>
<td>0.073</td>
<td></td>
</tr>
<tr>
<td>B3</td>
<td>0.082</td>
<td>0.077</td>
<td>0.089</td>
<td>0.103</td>
<td>0.090</td>
<td></td>
</tr>
<tr>
<td>B4</td>
<td>0.174</td>
<td>0.172</td>
<td>0.173</td>
<td>0.163</td>
<td>0.167</td>
<td></td>
</tr>
</tbody>
</table>
### 3.2 An Analysis of the Time Series Evolution of the Blue Carbon Economic Development in Daguang Bay, Jiangmen City

According to the blue carbon economic evaluation system of Daguang Bay in Jiangmen City constructed in this paper, the comprehensive score is calculated from the perspective of time series. The results are shown in Table 3. At the same time, in order to more intuitively show the development and changes of the blue carbon economy in the Daguang Bay of Jiangmen City The trend is described by a line chart, as shown in Figure 1.

**Table 3**: 2010-2019 Comprehensive scores of blue carbon economic development indicators in Daguang Bay, Jiangmen City

<table>
<thead>
<tr>
<th>City/District</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
<th>2014</th>
</tr>
</thead>
<tbody>
<tr>
<td>Xinhui District</td>
<td>0.691</td>
<td>0.682</td>
<td>0.685</td>
<td>0.669</td>
<td>0.657</td>
</tr>
<tr>
<td>Taishan City</td>
<td>0.500</td>
<td>0.513</td>
<td>0.486</td>
<td>0.482</td>
<td>0.516</td>
</tr>
<tr>
<td>Enping City</td>
<td>0.169</td>
<td>0.154</td>
<td>0.144</td>
<td>0.141</td>
<td>0.138</td>
</tr>
<tr>
<td>City/District</td>
<td>2015</td>
<td>2016</td>
<td>2017</td>
<td>2018</td>
<td>2019</td>
</tr>
<tr>
<td>Xinhui District</td>
<td>0.658</td>
<td>0.659</td>
<td>0.735</td>
<td>0.730</td>
<td>0.718</td>
</tr>
<tr>
<td>Taishan City</td>
<td>0.461</td>
<td>0.464</td>
<td>0.378</td>
<td>0.379</td>
<td>0.379</td>
</tr>
<tr>
<td>Enping City</td>
<td>0.172</td>
<td>0.197</td>
<td>0.235</td>
<td>0.274</td>
<td>0.288</td>
</tr>
</tbody>
</table>
Figure 1: The development and changes of the blue carbon economy in Daguang Bay, Jiangmen City

On the whole, the blue carbon economic development level of Xin hui District and Enping City, including Daguang Bay, Jiangmen City, from 2010 to 2019, showed an upward trend, while Taishan City as a whole was at a declining level, but Xin hui District and Taishan City The level of development of the blue carbon economy is generally better than that of Enping City. Before 2014, the development level of the blue carbon economy in Enping City was declining, and increased attention to it in 2014, after which the development level began to rise slowly. 2015 is a critical year for Xinhui District's industrial transformation and reform. The provinces and cities proposed an important deployment for building an advanced equipment manufacturing industry belt on the west bank of the Pearl River. The development of the Daguang Bay Economic Zone has been upgraded to a provincial strategy, and its development level has gradually risen in the past two years. The level of development has declined, and the development between regions is in a state of imbalance.

Therefore, intensifying the development of blue carbon economy in Taishan City and Enping City and achieving the balance of regional blue carbon economic development is of great significance for promoting the high-quality development of the marine economy of Daguang Bay.

4 CONCLUSIONS AND INSPIRATION

This paper analyzes the development situation from 2010 to 2019 by establishing a big data evaluation index system for the blue carbon economy of Jiangmen Daguang Bay, combined with the entropy method. The research conclusions are as follows:

- The impact of carbon sink on the development of the blue carbon economy of Daguang Bay in Jiangmen City is greater than that of the marine economy. The green coverage area and the added value of the tertiary industry are the main factors affecting the development of the blue carbon economy in Daguang Bay, Jiangmen City.

- From 2010 to 2019, the blue carbon economic development level of Daguang Bay of Jiangmen City, including Xinhui District and Enping City, showed an upward trend, while Taishan City was at a declining level, but the level of blue carbon economic development of
Xinhui District and Taishan City Overall, it is better than Enping City, and there is a phenomenon of uneven development between regions.

Through the above conclusions, in order to vigorously develop the blue carbon economy and realize the high-quality development of the marine economy of Daguang Bay. Get the following inspiration:

- Strengthen marine environmental protection. Strict environmental access mechanisms, keep the red line of ecological protection, strengthen the bottom line of environmental quality, and increase marine carbon sinks based on the scientific use of marine resources to effectively help achieve the goal of carbon neutrality.

- We will promote innovation in Marine science and technology. Innovation drives industrial development, and Marine scientific and technological innovation is an important strategic support for the high-quality development of China's Marine economy \(^{[7]}\). Combined with the current situation and potential of blue carbon development in Daguang Bay, support for blue carbon technology and industrial development research in the field of Marine related scientific research should be increased. At the same time, it will create a talent cultivation place for Daguang Bay, strengthen communication and cooperation with Wuyi University, Guangdong, Hong Kong and Macao higher education institutions, and plan to build higher education institutions and comprehensive colleges in Guanghai Bay to provide talent support for the economic development of the Daguang Bay Industrial Belt.

- Strengthen the blue carbon economic development of Taishan city and Enping City, narrow the gap with the new hui District, realize regional coordinated development, and jointly promote the high-quality development of Daguang Bay Marine economy.

Acknowledgements

This research was supported by the Guangdong Philosophy and Social Science Planning Project (Research on the trading mechanism of coastline resources in Guangdong - Hong Kong - Macao Greater Bay Area from the perspective of marine ecological civilization construction, Project No. GD20XGL14)

This research was supported by the Humanities and Social Sciences Research Planning Fund of the Ministry of Education (Research on dynamic pricing and trading mechanism of coastline resources in Guangdong-Hong Kong-Macao Greater Bay Area under the background of marine ecological civilization construction, Project No. 21YJA630097)

REFERENCES


Research on Evaluation of Innovation and Entrepreneurship Ability of Hebei Province Based on GPCA-HCA Model in Big Data Era

Shanshan Wang¹,a, Lihong Han¹,b*

¹075728389@qq.com, ¹070225171@qq.com
*Corresponding Author. Email: ¹070225171@qq.com

School of Economics and Management, Hebei University of Science and Technology, Shijiazhuang, China¹

Abstract: Big data is a new supporting force and driving force for innovation and entrepreneurship, which brings new opportunities to all subjects. Innovation and entrepreneurship under the background of big data is the new engine of regional economic development. Based on the panel data from 2016 to 2020, this paper constructs the evaluation system of innovation and entrepreneurship in eastern China by using the global principal component analysis method, and obtains the regional comprehensive ranking, and uses the comprehensive score to conduct hierarchical cluster analysis by using Ward algorithm to obtain the cluster pedigree map. The dynamic evaluation of Hebei province's innovation and entrepreneurship ability during the 13th Five-Year Plan period shows that the overall innovation and entrepreneurship ability of Hebei province is on the rise from 2016 to 2020, but the comprehensive ranking has been at a low level in the eastern region, and it is in the third echelon. The overall innovation and entrepreneurship ability is poor, the investment in innovation and entrepreneurship is insufficient, the output of innovation and entrepreneurship is low, and the basic environment of innovation and entrepreneurship needs to be improved.

Keywords: Big Data, Hebei Province, The 13th Five-Year Plan, Innovative and Entrepreneurial Ability, Global Principal Component Analysis, Hierarchical Cluster Analysis.

1 INTRODUCTION

After "mass entrepreneurship and innovation" was put forward in 2014, innovation and entrepreneurship became a new driving force to promote social and economic development in China. Under the background of big data, local governments, enterprises and science and education departments also paid more and more attention to big data, and they also increased their research investment in big data in innovation and entrepreneurship, such as increasing fixed assets investment in information transmission, software and information technology services. However, the development of innovation and entrepreneurship in different regions is not balanced, and the quality of innovation and entrepreneurship varies. Influenced by factors such as history and geographical location, Hebei's regional culture has no obvious characteristics. In recent years, Hebei Province has also vigorously advocated the development
of "mass entrepreneurship and innovation". For example, in 2020, Hebei Province put forward a number of policies and measures to promote the high-quality development of innovation and entrepreneurship. So, during the 13th Five-Year Plan period (2016-2020), what is the development level of innovation and entrepreneurship in Hebei Province? In this paper, the eastern region with similar geographical location and economic development level is selected as the research object. Based on the global principal component analysis and hierarchical cluster analysis, the innovation and entrepreneurship ability of eastern provinces (cities) from 2016 to 2020 is analyzed, and then the relative level of innovation and entrepreneurship ability of Hebei province and the gap between Hebei province and other eastern provinces (cities) are evaluated.

In recent years, many scholars have made research on the evaluation of regional innovation and entrepreneurship. HuangHuan and Ethan [2] sorted out the relationship between innovation and entrepreneurship, and then comprehensively evaluated the innovation and entrepreneurship ability of Sichuan Province in 2015 by using factor analysis. ParkZhefan [6] used TFP model and entropy weight -AHP model to evaluate the economic innovation and development capacity of Zhejiang Province from 2007 to 2017. HuPing [3] Using the weighted TOPSIS evaluation method to evaluate and analyze 30 provinces and cities in China in 2017, it is concluded that Zhejiang's comprehensive evaluation of innovation and entrepreneurship ranks fourth in the country.

Through combing the relevant literature, it is found that scholars have studied it from different angles and with different methods, which provides some ideas for this paper. However, most studies are based on a cross-sectional data, and data analysis cannot reflect the vertical development of innovation and entrepreneurship in a certain region. Based on this, this paper uses the collected panel data of eastern provinces (cities) in China from 2016 to 2020 to analyze the innovation and entrepreneurship capabilities of eastern provinces (cities) in China, and then evaluates the relative level of innovation and entrepreneurship capabilities in Hebei.

2 MATERIALS AND METHODS

2.1 Index Construction

Regional innovation and entrepreneurship capability is the result of the interaction of various innovation and entrepreneurship elements in a region, which can measure the efficiency of regional innovation and entrepreneurship. Based on the background of big data, this paper constructs an evaluation index system of regional innovation and entrepreneurship capability with dual input, dual output and dual environment as the main factors, as shown in Table 1.

<table>
<thead>
<tr>
<th>Main type</th>
<th>Index name</th>
<th>Index code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Investment</td>
<td>R&amp;D personne (human)</td>
<td>X1</td>
</tr>
<tr>
<td></td>
<td>Full-time equivalent of R&amp;D personnel (man-year)</td>
<td>X2</td>
</tr>
<tr>
<td></td>
<td>Science and technology expenditure (hundred million yuan)</td>
<td>X3</td>
</tr>
<tr>
<td></td>
<td>Internal expenditure of R&amp;D funds (ten thousand yuan)</td>
<td>X4</td>
</tr>
<tr>
<td></td>
<td>R&amp;D expenditure intensity (%)</td>
<td>X5</td>
</tr>
<tr>
<td>Output</td>
<td>Number of scientific papers published by institutions of higher learning (piece)</td>
<td>X6</td>
</tr>
</tbody>
</table>
2.2 Research methods

2.2.1 Gpca

In order to dynamically evaluate the innovation and entrepreneurship ability of Hebei Province and the gap between Hebei Province and other provinces (cities), this paper adopts global principal component analysis as the research method, which is abbreviated as GPCA. GPCA is based on the classic principal component analysis, adding time series, constructing three-dimensional data tables with the same index data by year, and replacing the original original data with comprehensive variables to dynamically analyze and study the research object. The specific steps are as follows:

(1) Establish a global data table. Suppose there are j indicators to be evaluated, which are recorded as $X_1$, $X_2$, $X_3$, ..., $X_j$. For n samples (this article refers to the provinces (cities) in the eastern part of China), there is a data table $X_t=(X_{ij})_{n \times j}$ in the t year, where $0 < t \leq T$. There are t data tables in a year. Arrange all the data tables before and after the year to form an $n \times t \times j$ three-dimensional time series data table, which is marked as: $X=(X_1, X_2, X_3, ..., X_t)$, so that each row in the matrix represents the data of the nth sample in a certain year.

(2) Data standardization. In order to make different types of index data comparable, the original data are uniformly transformed to eliminate the influence of dimensions. In this paper, Z-Score method is adopted for standardization. In this paper, the standard data table is denoted as $Z$.

$$Z_{ij} = \frac{X_{ij} - \mu_j}{\sigma_j} \quad (i = 1, 2, 3, ..., 10) \quad (1)$$

Among them, $Z_{ij}$ is the value of the jth index of the ith province (city) after dimensionless processing; $X_{ij}$ is the original data of the j index value of the ith province (city); $\mu_j$ is the average of the 10 provinces (cities) of the jth index data; $\sigma_j$ is the standard deviation of the jth index data.

(3) Data validity test. Verify the validity of the standardized data to determine whether the principal component analysis method can be used for these data. When $KMO \geq 0.6$, it means that...
the data is suitable for principal component analysis; If Bartlett's spherical test conclusion rejects the original hypothesis that variables are independent of each other, it means that variables are related, which is suitable for principal component analysis of data.

(4) Determine the principal component F. Calculate the covariance matrix R of the global data table X, and calculate the eigenvalues 1,2,3,...,p>0 of r, and the corresponding eigenvectors u1, u2, u3, up to calculate the variance contribution rate of the principal components:

$$a_k = \frac{\lambda_i}{\sum_{i=1}^{p} \lambda_i}$$

(2)

The cumulative variance contribution rate is:

$$a_1 + a_2 + \ldots + a_m = \frac{\sum_{i=1}^{m} \lambda_i}{\sum_{i=1}^{p} \lambda_i}$$

(3)

Select the principal components F1, F2, F3, …, Fm (0<m≤p) corresponding to the first m largest eigenvalues.

(5) Calculate the comprehensive score. Calculate the comprehensive score function of the research object through the scores of each principal component, as follows:

$$ZF = \sum_{i=1}^{m} \frac{\lambda_i}{q} F_i$$

(4)

Where i is the characteristic root of the ith principal component; q is the sum of the characteristic roots of the principal components, Fi is the standardized score of the ith principal component, and the principal component Fi is obtained by formula (5):

$$ZF = \sum_{i=1}^{m} \frac{\lambda_i}{q} F_i$$

(5)

Where m is the square root of the eigenvalue corresponding to the ith principal component divided by the ith column value in the matrix; Zm is the standardized data.

2.2.2 Hca

Hierarchical cluster analysis is a statistical method to classify data based on their own information. It divides data into several groups according to the distance (or similarity), and finally ensures that the differences within a group are as small as possible, while the differences between groups are as large as possible. This paper needs to evaluate the innovation and entrepreneurship ability of the provinces and cities in eastern China from 2016 to 2020. Therefore, three dimensions, sample number (N), indicator number (P) and time (T), need to be
considered in statistical analysis. Let $(X_1, X_2, \ldots, X_p)$ represent the characteristics of each sample, $(X_{ij}(t), t = 1, 2, \ldots, T)$ represent that the $j$ index of the $i$th sample is in the value of time $t$. When $X_{ij} > 0$, $i=1,2,\ldots,n$, $j=1,2,\ldots,p$, the Euclidean distance between sample $m$ and sample $n$ is defined as:

$$D_{mn} = \sqrt{\sum_{t=1}^{T} \sum_{j=1}^{p} [X_{mj}(t) - X_{nj}(t)]^2}$$ (6)

The distance function can be simplified as:

$$d_{mn} = \sqrt{\sum_{j=1}^{p} (X_{mj}^* - X_{nj}^*)^2}$$ (7)

### 2.3 Data Sources

In this paper, SPSS26.0 software is used to analyze and evaluate the innovation and entrepreneurship ability of scientific and technological talents in eastern China and the relative level of scientific and technological talents in Hebei province. The relevant index data of each province (city) from 2016 to 2020 comes from China Statistical Yearbook and China Science and Technology Statistical Yearbook.

### 3 RESULTS & DISCUSSION

#### 3.1 Empirical Process

The data of 10 provinces (cities) in the eastern region of China from 2016 to 2020 were arranged vertically by year to establish a time series global data table, and standardized processing was carried out. SPSS26.0 was used to measure the appropriateness of KMO sampling and Bartlett test to test the validity of the data. According to the data results in Table 2, the KMO test value is 0.825, which is greater than 0.5; Bartlett spherical test has a significance of 0.000, so the original hypothesis is rejected. (Luo, 2013)

<table>
<thead>
<tr>
<th>Table 2: Inspection of KMO and Bartlett</th>
</tr>
</thead>
<tbody>
<tr>
<td>Km sampling suitability quantity.</td>
</tr>
<tr>
<td>Bartlett's sphericity test</td>
</tr>
<tr>
<td>Approximate chi-square</td>
</tr>
<tr>
<td>Freedom</td>
</tr>
<tr>
<td>Significance</td>
</tr>
</tbody>
</table>
In order to transform many indexes into several comprehensive indexes and effectively reduce the dimension of data space, it is necessary to extract the principal components whose characteristic values are greater than 1, namely principal components F1, F2 and F3 (see Table 3). As can be seen from Table 3, the cumulative variance contribution rate of principal components F1, F2 and F3 reaches 84.567%. At the same time, the gravel diagram in Figure 1 shows that the characteristic values of the first three common factors change obviously, and the fourth characteristic value changes gradually. To sum up, these three principal components retain most of the information of the original data and can replace 17 original indicators [7].

### Table 3: Total variance of explanation

<table>
<thead>
<tr>
<th>Component part</th>
<th>Eigenvalue</th>
<th>Variance contribution rate%</th>
<th>Cumulative variance contribution rate%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9.728</td>
<td>57.226</td>
<td>57.226</td>
</tr>
<tr>
<td>2</td>
<td>3.157</td>
<td>18.571</td>
<td>75.797</td>
</tr>
<tr>
<td>3</td>
<td>1.491</td>
<td>8.769</td>
<td>84.567</td>
</tr>
</tbody>
</table>

The maximum variance orthogonal rotation method is used to rotate the initial load matrix by factors, and the rotation component matrix in Table 4 is obtained. According to the results in Table 4, it can be seen which principal component of the 17 indexes has higher factor load [7]. The first principal component (F1) has a high factor load on X1, X2, X3, X4, X6, X7, X8, X9, X10, X13, X15, X16, and it is named as the development factor of regional twinning ability. The second principal component (F2) has higher factor loads on X5, X11 and X14, so it is named as the regional double innovation input transformation capacity factor. The third principal component (F3) has a higher factor load on X12 and X17, so this principal component is named as the regional double-innovation economic development factor [1].

### Table 4: Rotation component matrix

<table>
<thead>
<tr>
<th>Ingredient</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td>0.956</td>
<td>0.150</td>
<td>0.204</td>
</tr>
<tr>
<td>X2</td>
<td>0.961</td>
<td>0.149</td>
<td>0.169</td>
</tr>
<tr>
<td>X3</td>
<td>0.917</td>
<td>0.284</td>
<td>-0.085</td>
</tr>
</tbody>
</table>
F1 = 0.3087Z1 + 0.3080Z2 + 0.2854Z3 + 0.2938Z4 + 0.0698Z5 + 0.0753Z11 + 0.2949Z13 - 0.0053Z14
F2 = -0.0022Z1 - 0.0019Z2 + 0.0879Z3 + 0.1510Z4 + 0.5383Z5 + 0.1097Z13 - 0.019Z14
F3 = 0.0081Z1 - 0.0251Z2 - 0.2464Z3 + 0.0739Z4 + 0.0318Z5 + 0.1119Z14 + 0.0100Z15 + 0.2014Z17

ZF = 0.6767F1 + 0.2196F2 + 0.1037F3

In the calculation formula of each common factor, the greater the coefficient of the index, the greater the promoting effect of the index on the common factor, and the smaller the opposite.

The comprehensive score expression is obtained from the principal component expression and the cumulative variance contribution rate of each principal component:

Table 5: Composition matrix

<table>
<thead>
<tr>
<th>Ingredient</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td>0.989</td>
<td>-0.004</td>
<td>0.009</td>
</tr>
<tr>
<td>X2</td>
<td>0.987</td>
<td>-0.003</td>
<td>-0.027</td>
</tr>
<tr>
<td>X3</td>
<td>0.914</td>
<td>0.152</td>
<td>-0.262</td>
</tr>
<tr>
<td>X4</td>
<td>0.941</td>
<td>0.261</td>
<td>0.079</td>
</tr>
<tr>
<td>X5</td>
<td>0.224</td>
<td>0.930</td>
<td>0.034</td>
</tr>
<tr>
<td>X6</td>
<td>0.442</td>
<td>0.301</td>
<td>0.069</td>
</tr>
<tr>
<td>X7</td>
<td>0.985</td>
<td>-0.065</td>
<td>-0.058</td>
</tr>
<tr>
<td>X8</td>
<td>0.957</td>
<td>-0.055</td>
<td>-0.127</td>
</tr>
<tr>
<td>X9</td>
<td>0.958</td>
<td>-0.210</td>
<td>-0.026</td>
</tr>
</tbody>
</table>
According to this, the scores, comprehensive scores and rankings of each principal component of innovation and entrepreneurship ability of 10 provinces (cities) in eastern China from 2016 to 2020 are calculated. The specific results are shown in Table 6.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.8</td>
<td>2.0</td>
<td>0.8</td>
<td>0.4</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>0.2</td>
<td>0.8</td>
<td>0.5</td>
<td>0.6</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>1.7</td>
<td>1.5</td>
<td>1.3</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>1.9</td>
<td>1.8</td>
<td>1.6</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>1.1</td>
<td>1.0</td>
<td>1.2</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>1.8</td>
<td>1.9</td>
<td>1.8</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>1.5</td>
<td>1.5</td>
<td>1.2</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>-</td>
<td>1.3</td>
<td>1.5</td>
<td>1.4</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>-</td>
<td>1.7</td>
<td>1.7</td>
<td>1.4</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>-</td>
<td>1.4</td>
<td>1.4</td>
<td>1.3</td>
<td>-</td>
</tr>
</tbody>
</table>

In this paper, the comprehensive scores of provinces calculated by the global principal component analysis method are used for cluster analysis, and SPSS26.0 clustering module is used for calculation. Q-type clustering in hierarchical clustering, Euclidean distance squared calculation and "Ward method" calculation are used for clustering method to obtain the cluster pedigree diagram of systematic cluster analysis, as shown in Figure 2. Considering the actual situation of the provinces and cities in eastern China, combined with the results of Figure 2, this
paper holds that the innovation and entrepreneurship capabilities of the provinces and cities in eastern China can be clearly divided into three groups, as shown in Table 7.

![Figure 2: Cluster pedigree diagram of innovation and entrepreneurship ability of provinces and cities in eastern China](image)

<table>
<thead>
<tr>
<th>Category</th>
<th>Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Guangdong, Jiangsu</td>
</tr>
<tr>
<td>2</td>
<td>Shandong, Zhejiang, Beijing</td>
</tr>
<tr>
<td>3</td>
<td>Hainan, Shanghai, Tianjin, Fujian, Hebei</td>
</tr>
</tbody>
</table>

### 3.2 Analysis of Empirical Results

#### 3.2.1 Analysis of Dynamic Characteristics of Innovation and Entrepreneurship in Eastern China

It can be seen from the comprehensive score and ranking table in Table 6 that the innovation and entrepreneurship capabilities of 10 provinces (cities) in the eastern region of China are generally on the rise from 2016 to 2020, but the innovation and entrepreneurship capabilities of different regions are quite different, and the innovation and entrepreneurship capabilities of the same region are also fluctuating in different years. According to the cluster pedigree diagram obtained by cluster analysis, the innovation and entrepreneurship capabilities of the eastern provinces and cities of China can be divided into the following three categories.: the first category includes Guangdong and Jiangsu. In 2016-2020, Guangdong and Jiangsu had strong innovation and entrepreneurship ability, and their comprehensive scores were much higher than those of other provinces (cities). Among them, Guangdong ranked first in 2016-2020, and its comprehensive scores rose the fastest, from 2.00 in 2016 to 5.79 in 2020; Jiangsu ranked second in 2016-2020, rising from 1.84 in 2016 to 4.16 in 2020; The second category includes Beijing, Zhejiang and Shandong. These three provinces and cities ranked 3rd-5th, with Beijing’s comprehensive score of 0.15, Shandong's comprehensive score of -0.18 and Zhejiang's comprehensive score of -0.21 in 2016. In 2017, the comprehensive scores of the three provinces...
and cities were all positive and developed rapidly. The third category includes Shanghai, Hebei, Fujian, Tianjin and Hainan. The comprehensive scores of the third-class provinces and cities in 2016-2020 were all negative, and compared with the first two types of provinces and cities, the development of innovation and entrepreneurship in Hainan was the slowest, ranking 10th in 2016-2020.

There is polarization in the comprehensive scores among provinces and cities in the eastern region. For example, in 2020, the average comprehensive score of innovation and entrepreneurship in 10 provinces (cities) in the eastern region was 0.704, that of Guangdong province was 5.79, and that of Hainan was -3.12, with a difference of 8.91 between them. The polarization phenomenon is serious.

3.2.2 Analysis of Dynamic Characteristics of Innovation and Entrepreneurship Ability of Scientific and Technological Talents in Hebei Province

The variance contribution rate of the development factors of regional innovation capability in first principal component reached 57.226%, which had the greatest impact on regional innovation and entrepreneurship. As can be seen from Figure 2, the development factor of regional dual-innovation capability in Hebei Province has developed fastest, rising from -2.73 in 2016 to -1.34 in 2020. It can also be seen from Table 8 that the ranking of regional dual-innovation capability development factor has risen from 9th in 2016 to 7th in 2020. The increase of R&D personnel, the full-time equivalent of R&D personnel, the expenditure of science and technology and the internal expenditure of R&D funds in Hebei Province, as well as the emphasis on supporting infrastructure of education, innovation and entrepreneurship have played a key role in the promotion of the ranking of regional innovation capability development factors in Hebei Province. From 2016 to 2020, the internal expenditure of R&D funds in Hebei Province increased from 38.343 billion yuan to 63.437 billion yuan, an increase of 65.45%; although the number of R&D personnel, the full-time equivalent of R&D personnel and the expenditure on science and technology fluctuate slightly, the overall situation is on the rise. From 2016 to 2020, Hebei Province's education funds ranked fifth among all provinces and cities in the eastern region, which provided good educational support for innovation and entrepreneurship in Hebei Province. The number of scientific and technological enterprises in Hebei Province has also increased year by year, from 102 in 2016 to 282 in 2020, which has improved the score of regional innovation capability development factor in Hebei Province. In recent years, Hebei Province has also continuously increased its support for big data and increased its investment in fixed assets of information transmission, software and information technology services. In 2020, the fixed assets investment of information transmission, software and information technology services in Hebei Province was 56.144 billion yuan, ranking third in the eastern region of China. The backward economic output of innovation and entrepreneurship and the output of knowledge and technology in Hebei Province are the main reasons for the backward ranking of regional innovation capability development factors in Hebei Province. Compared with other provinces and cities in the east of China, Hebei has fewer coastal areas, and its market development and openness are lower than those of other coastal provinces and cities in the east. In 2016, the sales revenue of new products of industrial enterprises above designated size in Hebei reached 392.314 billion yuan, and in 2020, the sales revenue reached 719.098 billion yuan, with a rapid growth, ranking 6th-7th in 2016-2020. In the output of knowledge and technology, the number of papers published by universities in Hebei province,
the number of domestic patent applications accepted and the number of domestic patent applications authorized are in a inferior position, lagging behind other provinces and cities in eastern China, ranking 8th-9th.

![Figure 3: Chart of principal components of innovation and entrepreneurship ability in Hebei Province from 2016 to 2020](image)

The variance contribution rate of the second principal component is 18.571%. As can be seen from Figure 3, the development of regional dual innovation transformation capacity factor in Hebei Province is relatively slow. In 2016 and 2017, the score of this component was -1.51. In 2018, the score slightly increased to -1.42, then decreased to -1.44 in 2019, and rose to -1.31 in 2020. The ranking of this factor is also at the bottom, 9-10, which is mainly due to the fact that Hebei Province's R&D investment intensity, technology market turnover and the average number of college students per 100,000 population are lower than other eastern provinces and cities. From 2016 to 2020, Hebei Province ranked 8th in terms of technology turnover, which is far from the developed cities in the east.

<table>
<thead>
<tr>
<th></th>
<th>2016</th>
<th>2017</th>
<th>2018</th>
<th>2019</th>
<th>2020</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>-2.73</td>
<td>9</td>
<td>-2.45</td>
<td>9</td>
<td>-2.26</td>
</tr>
<tr>
<td>F2</td>
<td>-1.51</td>
<td>10</td>
<td>-1.51</td>
<td>9</td>
<td>-1.42</td>
</tr>
<tr>
<td>F3</td>
<td>0.02</td>
<td>4</td>
<td>0.16</td>
<td>8</td>
<td>0.33</td>
</tr>
</tbody>
</table>

The third principal component, the regional double-innovation economic development factor, was on the rise in 2016-2019, from 0.02 in 2016 to 0.77 in 2019, and declined slightly to 0.70 in 2020, but the ranking fluctuated greatly. In 2016, the per capita GDP of Hebei was 43,062 yuan, ranking 9th in the eastern region, which was relatively backward. However, in 2016, the investment in fixed assets of scientific research and technical services in Hebei was 35.18 billion yuan, ranking third, ranking first. Therefore, comprehensively speaking, the regional dual-innovation economic development factors ranked first in 2016, and the variance contribution rate of the regional dual-innovation economic development factors was 8.769%, which had a limited impact on the innovation and entrepreneurship of Hebei.

From the cluster genealogy, we can know that Hebei province is in the third group, and its comprehensive ability of innovation and entrepreneurship is weak, and the comprehensive scores of Hebei from 2016 to 2020 obtained by the global principal component analysis method are all negative, with the comprehensive score of -2.18 in 2016 and -1.12 in 2020. In terms of
ranking, the comprehensive ranking of innovation and entrepreneurship in Hebei province from 2016 to 2018 ranks ninth. In 2019-2020, it ranked seventh. It can be seen that Hebei's innovation and entrepreneurship ability is gradually improving, but it is still lower than the average level in the eastern region, and the gap between Hebei and the average level in the eastern region is expanding.

4 CONCLUSIONS AND SUGGESTIONS

By comparing with other eastern provinces (cities), this paper analyzes and evaluates the innovation and entrepreneurship ability of Hebei. The empirical results show that from 2016 to 2020, the innovation and entrepreneurship capacity of Hebei is on the rise, but Hebei is located in the third echelon in the eastern region, and there is a big gap with other developed provinces (cities) in the east, which is characterized by poor innovation and entrepreneurship capacity, insufficient investment in innovation and entrepreneurship, low innovation and entrepreneurship output, and the basic environment for innovation and entrepreneurship needs to be improved. According to the above evaluation results and the present situation of innovation and entrepreneurship development in Hebei, the following suggestions are put forward.

First of all, all market players should conform to the development requirements of the era of big data, actively integrate into innovation and entrepreneurship, and improve the quality of innovation and entrepreneurship. Secondly, Hebei Province should establish an innovation and entrepreneurship data platform based on big data, and use information technology and computer technology to form an innovation and entrepreneurship atmosphere of data analysis, management and decision-making, so as to improve the efficiency of innovation and entrepreneurship. At the same time, Hebei should change from government-led to market-led in technological innovation, and constantly improve the construction of technology market system. Finally, Hebei should improve the regional coverage of new infrastructure, make rational layout, make overall planning and develop key areas and industries.

Acknowledgements: In 2021, the key research project of humanities and social sciences in colleges and universities of Hebei Province, Research on the Implementation Path of xiong'an new area Innovation and Entrepreneurship Elements Agglomeration (SD 2021021)

REFERENCES


Chao Li
624299703@qq.com
School of accounting, Wuhan Qingchuan University, Hongshan District, Wuhan City, Hubei Province, China

Abstract: With the continuous progress of modern science and technology, the arrival of the era of big data, if the advantages of the era of big data can be applied to financial management, it will create more development opportunities for enterprises. Based on this, this paper collects the required data through the data acquisition algorithm model, and mainly discusses the innovation strategy of enterprise financial management in the era of big data. The paper points out that under the background of big data, 1) Enterprises lack of advanced financial management concepts; 2) Enterprises lack the infrastructure to use big data technology; 3) The quality of enterprise financial management personnel is uneven. To solve these problems, this paper puts forward the following suggestions: 1) Establish advanced financial management concept; 2) Introducing sound infrastructure for the application of big data technologies; 3) Strengthen the education and training of enterprise financial management personnel, improve the comprehensive quality of staff. This paper hopes that these suggestions can provide reference for the managers of financial departments of enterprises.

Keywords: The Era of Big Data, Financial Management Activities, Data Collection Algorithm Model, Problems and Suggestions.

1 INTRODUCTION

Big data involves a wide range of fields, including astronomy, atmospheric science, genomics and other fields, as well as military reconnaissance, financial big data, medical big data and other industries [5, 8, 9]. At present, the advantages of big data are gradually highlighted, and the industries penetrated by big data are constantly expanding, such as social networking [6], e-commerce, communication record details and other industries developed by big data help enterprises continue to expand new businesses and innovate operation forms [4, 10]. Since the introduction of the concept of "big data", enterprises have been comprehensively improved and perfected in the estimation of product sales quantity, behaviour evaluation of mass consumption, accurate judgment of marketing scope and supply of stored goods [2].

In the era of big data, there are three changes in public perception: big data is the whole data, not random extraction (11). Big data is a general direction, not an accurate guide. Big data is correlation, not cause-effect correlation. The application and development of big data
improves the effectiveness of data for development, provides accurate market data for the judgment of enterprise management departments, and improves the quality of enterprise decision-making [3].

Under the background of the big data era, the vast majority of enterprises determine profit acquisition as the direction of operation [1, 7]. Therefore, financial management plays a crucial role in enterprise management activities. In this case, it is of important theoretical and practical significance to explore the innovative strategies of enterprise financial management against the background of the big data era to promote social development and progress. What is the impact of the era of big data on the financial management of enterprises and how to improve the financial management of enterprises under the background of the era of big data are urgent problems to be solved at present.

2 DATA COLLECTION ALGORITHM MODEL

Due to the need for a wide range of data, so this paper with the aid of quantum genetic algorithm to collect data. The flow of this algorithm is as follows:

Begin t=0
(1) initialize Q(t0)
(2) make P(t) by observing the states if Q(t)
(3) evaluate P(t)
(4) store the best solution among P(t) into B(t)
   While (not termination condition) do
      Begin t=t+1
      (5) make P(t) by observing the states if Q(t-1)
      (6) evaluate P(t)
      (7) update Q(t) using Q-gates
      (8) store the best solution among B(t-1) and P(t) into B(t)
      (9) store the best solution b among B(t)
      (10) if (migrate condition)
           Then migrate b or to B(t) globally or locally
      END
   END

The algorithm first generates the initial population Q(t0). Where, the gene location (α_i, β_i) of the population is initialized as (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}), so that the population chromosome is equal to the superposition of all states and equal probabilities:

\[ \psi_{q_i} &= \sum_{k=1}^{2^m} \frac{1}{\sqrt{2^m}} S_k \]

where, S is composed of binary string (x_1, x_2, ..., x_m) with length m, representing the kth state of the chromosome. x_i (i=1,2,3,...,m) is either 0 or 1.
The second step of the algorithm is to scan all the individuals in the population once and obtain a set of solutions \( p(\ell) = \{p_1^{\ell}, p_2^{\ell}, \ldots, p_k^{\ell}\} \). Here, \( p \) is used to represent the \( J \)th solution in the \( T \)-generation population, which is composed of \( m \) bits of binary, where each bit is 0 or 1, and the bit value of binary is the probability amplitude of quantum bits \( (|\alpha_i|^2 \text{ or } |\beta_i|^2, i = 1, 2, \ldots, m) \) decided.

Then determine whether the loop termination condition is satisfied. If yes, the algorithm ends. Otherwise, the relevant quantum genetic manipulation is performed to further manipulate the chromosomes and create a new population. As the loop continues to execute, the algorithm will slowly converge to the optimal solution.

3 THE NECESSITY OF APPLYING BIG DATA TECHNOLOGY IN THE FIELD OF ENTERPRISE FINANCIAL MANAGEMENT

3.1 The Application of Big Data Technology is an Inevitable Trend of Historical Development

The application of big data technology can fundamentally improve the efficiency of enterprise financial management, significantly improve the quality of enterprise follow-up work, and make enterprise operation more coordinated and unified. Big data technology enables departments to perform tasks more systematically. Under the guidance of big data technology, all kinds of work can be operated methodically, so that the management level of enterprises can be continuously improved imperceptibly. The popularization and wide application of big data technology is the inevitable trend of historical development. More and more enterprises actively use it in daily operation. Therefore, the financial staff of each enterprise should set up advanced ideas to keep pace with The Times and keep up with the development trend of The Times, so that the implementation of more work of the enterprise has intelligent characteristics.

The innovation of data management system can improve the implementation efficiency of subsequent financial work of enterprises. The managers involved need to take a longer view. Select and apply the most suitable data and information integration software according to the actual situation of the enterprise, so as to improve the quality and efficiency of the enterprise financial management. Enterprise managers can make a more suitable enterprise operation plan according to the information displayed in relevant software, and omit redundant steps of enterprise management. Effectively improve the quality of related work, keep pace with The Times, so that the enterprise financial management can be guided by new technology with modern characteristics, so that the enterprise internal personnel can understand the market development situation, to ensure that the quality of the work and the decision can be significantly improved.

3.2 The Application of Big Data Technology Can Effectively Meet the Actual Development Needs of Enterprises

If big data technology is widely used in enterprise financial management, enterprise managers need to have a long-term vision of development. The advanced information technology and
the actual situation of the enterprise organic integration. Fundamentally change the original pattern of traditional financial management. The implementation of each financial management work can grasp the essence of innovative thinking. Promote the implementation of follow-up work to break the shackles of old ideas, and fundamentally improve the efficiency of financial management work. Enterprises actively use big data technology to perform various financial management tasks, in fact, is an important reflection of self-improvement. Only in this way, enterprise managers can develop more perfect enterprise management policies, so that the implementation of all work can be based on evidence. Therefore, it is extremely necessary to make reasonable planning for the work in the field of financial management. This requires workers in the field to have sufficient sense of innovation, so that enterprises can take advantage of the complex market environment. The gradual popularization of big data technology can effectively help enterprises expand development space, promote enterprises to achieve higher achievements in financial management, and obtain more customer information resources. The application of big data technology also makes more employees realize the importance of innovation consciousness, and they can actively use their spare time to study and learn advanced technologies, which promotes the comprehensive ability of each employee to be significantly improved.

3.3 The Automated Office Model is Convenient to Implement

With the rapid development of big data technology in recent years, the speed of information transmission has been improved. The speed of information transmission has been significantly improved, and the efficiency of financial management has also been improved. In the past financial management of enterprises, many tasks need to be completed manually, which leads to the implementation of a lot of work prone to manual errors. For each job management will face more complex problems, resulting in the quality of financial management work is seriously affected, the value of management work is greatly discounted. A lot of important data information is likely to be lost or omitted in the manual operation, which will give the enterprise financial management work buried security risks. The active application of big data technology is the inevitable trend of historical development, and also the problem that enterprises need to focus on in the process of operation. The application of big data technology can improve the efficiency of the original complex financial data analysis work, so that the staff of various departments of the enterprise can have a deep understanding of the actual business status of the enterprise, and can promote the staff of the enterprise to understand the key content of the follow-up work in the more detailed financial data statements, and determine the work objectives in time.

4 ANALYSES OF THE QUESTIONS OF ENTERPRISE FINANCIAL MANAGEMENT IN THE ERA OF BIG DATA

4.1 Enterprises Lack of Advanced Financial Management Concepts

Now, the era of big data has come, and many enterprises have noticed the impact of the era of big data on enterprises. However, some enterprises in the process of carrying out financial management activities, still use the previous management forms and management ideas, and have not made corresponding changes. In a word, the form and concept of financial
management of enterprises lag behind, mainly in the following three aspects. First of all, enterprises do not take big data technology as an effective measure and way to carry out financial management activities. When enterprises carry out financial management activities, they still use the previous technical analysis methods, resulting in inaccurate data analysis results. Secondly, the enterprise financial management personnel did not transmit or convey the big data application requirements to the enterprise's internal financial workers. This leads to enterprise financial workers in the process of carrying out financial management activities, there is no awareness of the application of big data analysis technology. Finally, many enterprise executives believe that the cost of applying big data technology is too high and the benefits obtained are too little, so they do not support the use of big data technology to analyse data within the enterprise. It is precisely because of such backward financial management concept, hindering the progress of enterprises, resulting in many enterprises cannot feel the benefits of big data technology.

4.2 Enterprises Lack the Infrastructure to Use Big Data Technology

In the process of applying big data technology, many enterprises are faced with a common problem, which is the lack of reasonable use of big data technology infrastructure. The so-called big data means that the amount of data is extremely large. The computers purchased by enterprises in the past simply cannot meet the requirements of data storage in the era of big data. This is because the capacity of the traditional computer is too small, the amount of data of the enterprise is too large, there is no way to meet the needs of the enterprise, resulting in the enterprise cannot use the corresponding technology to implement big data analysis. Not only that, many small and medium-sized enterprises have not built the corresponding financial management processing system, which more limited the application of big data technology. If enterprises want to make use of big data technology in the process of carrying out financial management activities, they must build sound infrastructure that is compatible with the use of big data technology, and lay a solid foundation for enterprises to apply the corresponding big data technology.

4.3 The Quality of Enterprise Financial Management Personnel is Uneven

With the advent of the era of big data, enterprises have gradually improved their standards for financial management personnel. The standard of this quality is not only reflected in meeting the requirements of professional literacy, but also embodied in information technology literacy. One of the most critical is the knowledge related to the computer level. Because enterprise managers must carry out computer-related operations in the process of using big data technology to solve financial activities. If employees do not understand the basic computer control technology, there is no way to complete the corresponding work. However, from the current situation of comprehensive quality of most enterprise workers, many financial managers lack the skills to apply big data technology; The quality of enterprise financial management workers is uneven; As a result, enterprises cannot reasonably apply big data skills in the process of financial management activities. Therefore, enterprises should pay attention to the training and selection of enterprise financial management workers, improve the comprehensive quality of enterprise financial management workers, and lay a foundation for enterprise personnel to apply big data technology to carry out financial management work.
5 NEW STRATEGIES FOR ENTERPRISE FINANCIAL MANAGEMENT IN THE ERA OF BIG DATA

5.1 Establish Advanced Financial Management Concept

Enterprises must change the concept of financial management, establish a scientific concept of financial management that conforms to the development needs of the era of big data, and create a good environment for the application of big data skills in financial management. Enterprises should guide employees to use big data skills to carry out financial management activities and establish financial management concepts. Enterprise managers or senior executives need to deeply explore knowledge related to big data and guide employees to establish a sense of big data application. Enterprise managers should actively create an enterprise culture that is helpful to the application of big data, and encourage and support enterprise financial management workers to actively practice big data analysis and other skills.

At the same time, the enterprise should regularly carry out publicity and training activities related to the era of big data, so that the enterprise financial management workers abandon the previous traditional backward financial ideas, and establish the big data thought system of financial management personnel. The enterprise should change the salary and reward system and associate the salary and reward system with the application of big data technology, so that the staff should not only do things well, but also pay attention to the efficiency of work, so as to ensure the real application of big data technology in the process of financial management activities of the enterprise, so that it can play an important role.

Figure 1 shows the analysis of the profit level of Chinese financial companies from 2014 to 2019, suggesting that financial management occupies an increasing proportion in the Chinese market and is becoming more and more important.

![Figure 1: The analysis of the profit level of Chinese financial companies from 2014 to 2019 (unit: 100 million yuan).](image)

5.2 Introducing Sound Infrastructure for the Application of Big Data Technologies

If enterprises want to reasonably apply the corresponding big data technology, they should first improve the basic equipment. Only when the hardware is in place are employees likely to discuss how to apply big data technology. However, from the current actual situation, the vast majority of Chinese enterprises have not introduced sound infrastructure to assist enterprises to apply big data technology. Many companies still use old computers for data analysis, which
constrains the use of big data within the enterprise. If enterprises want to increase the frequency of big data application, they need to check whether they have the basic equipment related to big data application. According to the verification results, the corresponding facilities of the enterprise shall be improved or replaced. Generally speaking, this process is very easy to suffer resistance from the corresponding staff, especially the company's investment staff. Because investors will feel that this will increase the cost load of the enterprise. It is not helpful to the development of the enterprise. At this time, the actual management of the enterprise is required to communicate and adjust, to ensure that the enterprise can successfully complete the introduction of basic equipment. The actual management personnel of enterprises must communicate with investment personnel in the form of teaching, and cannot use too extreme ways. For example, coerce investment personnel, require investment personnel must accept their own planning scheme. Such behaviour will only lead to a dilemma in communication, unable to give full play to the role of big data technology.

5.3 Strengthen the Education and Training of Enterprise Financial Management Personnel, Improve the Comprehensive Quality of Staff

When enterprises use big data technology for financial management, they also need to deal with an important problem, that is, the quality of personnel. If the enterprise lacks some highly literate financial management workers, then in the process of carrying out financial management activities, the application of big data will only become an empty word and cannot be implemented. Therefore, enterprises must enhance the training of enterprise financial management staff. More staff training is needed. Not only should the corresponding professional quality of financial management personnel be improved, but also the application level of big data technology of financial management personnel should be further improved. Efficient control of the computer, for the enterprise financial management work efficient development to provide talent support.

6 CONCLUSIONS

To sum up, enterprises must learn how to apply big data technology in financial management under the background of big data era, effectively improve the application ability and application quality of big data technology of enterprises, improve the level of enterprise financial management, and promote the sustainable development of enterprises.

REFERENCES

Big Data Evaluation on Marine Ecological Civilization in Guangdong Province

Jianhua Xiao¹, Qiqin Hu¹*

*Corresponding author and e-mail: Qiqin Hu, h15083831728@163.com

School of Economics and Management, Wuyi University, Jiangmen, China¹

Abstract: In this paper, entropy method and coupling coordination degree model were used to evaluate the level of marine ecological civilization construction in Guangdong Province from 2009 to 2018. Fourteen indicators were selected from four aspects, including economy, society, ecology, environment and ecology. A performance evaluation system was established for marine ecological civilization construction in Guangdong Province. The results show that the comprehensive development index of society and the environment is low. Thus, in the future construction process, we need to pay more attention to the quality of life and strive to improve living standards. At the same time, we need to strengthen the protection and governance of the environment. The coupling coordination of the performance evaluation system of marine ecological civilization shows a slow growth, which transitions from the basic coordination stage to the good coordination stage. This indicates that the construction level has been improved but still needs to be strengthened.

Keywords: Marine Ecological Civilization, Entropy Value Method, Coupling Coordination Degree Model, Guangdong Province.

1 INTRODUCTION

In 2019, the fourth Plenary Session of the 19th CPC Central Committee required improving the exploitation and protection of the marine resources system and strengthening the protection of Marine natural ecology from the perspective of "adhering to and improving the system of ecological civilization" [1]. Marine ecological civilization construction is an important part of ecological civilization construction and coastal economy. According to preliminary calculation, China's Marine GDP in 2019 was 8941.5 billion yuan with an increase of 6.2% over the previous year, and the proportion of marine industry in GDP was 9.0%[2]. While providing various means of production, the ocean is also facing increasingly serious environmental problems [3,4]. In recent years, Chinese scholars have mainly studied the connotation definition, evaluation system construction and practice of marine ecological civilization construction.

In terms of the connotation of marine ecological civilization construction, Zhu(2017)[5] believes that the connotation includes three aspects. First, consciousness, civilized behavior, and systematic civilization of marine ecological civilization. Second, marine ecological civilization behavior. Third, marine ecological system civilization. In terms of the construction of its evaluation system, Feng and Ye (2021)[6] studied the construction of marine ecological civilization in Zhejiang Province and selected 34 indicators to construct an evaluation system...
from six aspects of community, economy, governance, culture, resources, and ecology to analyze
the construction level of Zhejiang Province. In terms of practice, Gao(2016)[7] studied the
advantages, existing problems, and future work direction of Shengsi City’s marine ecological
civilization construction and put forward suggestions for the existing problems. This is conducive
to the development of Shengsi City and provides a reference for other coastal cities.

2 DATA SOURCES AND RESEARCH METHODS

For Guangdong Province, we propose a set of objectives and a comprehensive performance
evaluation index system from the perspectives of economy, society, environment and science,
and technology to supplement the existing achievements and provide relevant basis for the
construction of other regions.

2.1 Evaluation Index System of Marine Ecological Civilization Development

Based on the research of Feng and Ye(2021)[6], Miao and Wang(2020)[8], and the principles of
quantification, timeliness, representativeness and feasibility, we constructed an evaluation index
system (Table 1) comprising target layer, criterion layer and index layer. The construction
process was combined with the actual situation to be representative and feasible.

2.2 Data Sources

In order to explore and evaluate the development of Marine ecological civilization in Guangdong
Province, economic, social, resource, scientific and technological data were collected from China
Marine Statistical Yearbook, China Statistical Yearbook, Guangdong Statistical Yearbook, as
well as statistical bulletins and work reports of relevant departments of Guangdong Province
from 2009 to 2018. The study covers Guangdong Province. Individual missing data are filled in
with exponential smoothing method on the basis of previous years’ data.

Table 1: TABLE I. EVALUATION INDEX SYSTEM.

<table>
<thead>
<tr>
<th>The target layer</th>
<th>Rule layer</th>
<th>Index layer</th>
<th>Index of the unit</th>
<th>Index attribute</th>
<th>The weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance level of Marine ecological civilization construction in Guangdong Province</td>
<td>economic (0.2709)</td>
<td>Gross Marine product</td>
<td>One hundred million yuan</td>
<td>+</td>
<td>0.067</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Proportion of GROSS Marine product in gross regional product</td>
<td>%</td>
<td>+</td>
<td>0.0633</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The proportion of Marine tertiary industry in gross Marine product</td>
<td>%</td>
<td>+</td>
<td>0.0819</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Direct economic losses caused by Marine disasters</td>
<td>One hundred million yuan</td>
<td>-</td>
<td>0.0587</td>
</tr>
<tr>
<td></td>
<td>social</td>
<td>GDP per capita</td>
<td>yuan</td>
<td>+</td>
<td>0.0666</td>
</tr>
</tbody>
</table>
2.3 Index Weight of Evaluation System

The weight of indicators can reflect the importance of indicators to the construction of ecological civilization. In order to avoid the influence caused by subjective factors, objective weighting method is adopted in this paper. The larger $X_{ij}$ is, the better the evaluation is, which indicates that $X_{ij}$ is a positive indicator, and vice versa. In order to eliminate the influence of different dimensions and orders of evaluation indicators, the evaluation indicators need to be standardized before analysis. After standardizing, the part of the result is 0, which makes the subsequent logarithmic calculation meaningless. Therefore, a positive number slightly greater than zero is added to the result of all calculations.

Standardizing is carried out by following equations.

\[
X_{ij}^* = \frac{X_{ij} - X_{j\min}}{X_{j\max} - X_{j\min}} + 0.0001
\]  \hspace{1cm} (1)

\[
X_{ij}^* = \frac{X_{j\max} - X_{ij}}{X_{j\max} - X_{j\min}} + 0.0001
\]  \hspace{1cm} (2)

<table>
<thead>
<tr>
<th>The target layer</th>
<th>Rule layer</th>
<th>Index layer</th>
<th>Index of the unit</th>
<th>Index attribute</th>
<th>The weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.1799)</td>
<td></td>
<td>Number of employed persons</td>
<td>Ten thousand people</td>
<td>+</td>
<td>0.0591</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Proportion of urban population</td>
<td>%</td>
<td>+</td>
<td>0.0542</td>
</tr>
<tr>
<td>The environment (0.2255)</td>
<td></td>
<td>Proportion of Marine area of first and second class water quality in nearshore area</td>
<td>%</td>
<td>+</td>
<td>0.0532</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Investment in environmental pollution control as a proportion of GDP</td>
<td>%</td>
<td>+</td>
<td>0.0775</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Discharge of sewage from the sea</td>
<td>One hundred million tons of</td>
<td>-</td>
<td>0.0948</td>
</tr>
<tr>
<td>Science and technology (0.3237)</td>
<td></td>
<td>Number of Marine scientific research institutions</td>
<td>a</td>
<td>+</td>
<td>0.0702</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Number of Marine science and technology patents granted</td>
<td>item</td>
<td>+</td>
<td>0.1014</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Marine science and Technology activists</td>
<td>people</td>
<td>+</td>
<td>0.0863</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The number of graduates with master's degree and doctor's degree in oceanography</td>
<td>people</td>
<td>+</td>
<td>0.0658</td>
</tr>
</tbody>
</table>
The proportion of $J$ in $I$ is

$$Y_{ij} = \frac{x_{ij}^{*}}{\sum_{j=1}^{n}x_{ij}^{*}} \quad (3)$$

The following equation is used to calculate the information entropy of $j$.

$$e_j = -k \sum_{i=1}^{n}(Y_{ij} \times \ln(Y_{ij})) \quad (4)$$

Information entropy was used to calculate the different coefficients of $j$.

$$d_j = 1 - e_j \quad (5)$$

The difference coefficient is used to calculate the weight of $j$.

$$W_j = \frac{d_j}{\sum_{i=1}^{n}d_j} \quad (6)$$

### 2.4 Coupling Coordination Degree Model

After determining the index weight, calculate the comprehensive score of the ecological civilization index:

$$D_i = \sum_{j=1}^{n}(W_j \times X_{ij}^{*}) \quad (7)$$

$D_i$ is the comprehensive evaluation score, and $n$ is the number of indicators.

### 2.5 Coupling Coordination Degree Model

The coupling degree means the degree of interaction between two or more systems. The smaller value indicates that the system develops in the direction of disorder and conversely in the direction of order. The degree of coupling among economic, social, environmental and technological subsystems reflects the degree of interaction among each system. Since the number of subsystems is four, the equation is as follows.

$$C = 4 \times \left[ \frac{U_1 \times U_2 \times U_3 \times U_4}{(U_1 + U_2 + U_3 + U_4)^{\frac{1}{2}}} \right]^{\frac{1}{2}}$$

where $C$ represents the coupling degree and ranges from 0 to 1, $U_1$, $U_2$, $U_3$, and $U_4$ represent the comprehensive development indicators of the four subsystems: economy, society, environment and science and technology. In this paper, the coupling degree is divided into four sub-values: 0–0.30, 0.30–0.50, 0.50–0.80, and 0.80–1, which respectively represent four stages: low level coupling, antagonism, run-in, and high-level coupling[9].
The indexes in the evaluation system are dynamic and unbalanced, and errors may be generated only depending on the coupling degree. Therefore, to avoid such errors, the following coupling coordination function is used.

\[ D = \sqrt{C \times T} \] (9)

\[ T = aU_1 + bU_2 + cU_3 + dU_4 \] (10)

where \( D \) is the coupling coordination degree, \( T \) is the comprehensive coordination index of economy, society, environment and science and technology subsystems, \( A, B, C, \) and \( D \) are undetermined parameters, and \( A + B + C + D = 1 \). Since the coordination effects of the four subsystems are the same, \( 1/4 \) of \( A, B, C, \) and \( D \) are selected in this paper. Based on relevant research results\cite{10,11} and the actual situation, we divided the coupling coordination degree into the following five types. When \( D \) is greater than 0 and less than or equal to 0.1, it represents serious disorder. When \( D \) is greater than 0.1 and less than or equal to 0.2, it is moderate disorder. When \( D \) is greater than 0.2 and less than or equal to 0.3, it represents basic coordination. When \( D \) is greater than 0.3 and less than or equal to 0.4, it indicates intermediate coordination. Good coordination is indicated when \( D \) is greater than 0.4 and equal to 0.7. When \( D \) is greater than 0.7 and equal to or equal to 1, the coordination is good.

3 RESULTS AND ANALYSIS

3.1 Comprehensive Evaluation Index Analysis

According to the real data of Guangdong province from 2009 to 2018 and Eqs. (1)-(7), the comprehensive development index of the four subsystems of economy, society, environment and science and technology is calculated as shown in Table 2 and Figure 1.

<table>
<thead>
<tr>
<th>Year</th>
<th>Economics</th>
<th>Society</th>
<th>The environment</th>
<th>Science and technology</th>
<th>Construction level of marine ecological civilization</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009</td>
<td>0.04286</td>
<td>0.00002</td>
<td>0.11724</td>
<td>0.09327</td>
<td>0.25338</td>
</tr>
<tr>
<td>2010</td>
<td>0.06421</td>
<td>0.03739</td>
<td>0.14034</td>
<td>0.07410</td>
<td>0.31603</td>
</tr>
<tr>
<td>2011</td>
<td>0.07413</td>
<td>0.05296</td>
<td>0.11701</td>
<td>0.08411</td>
<td>0.32821</td>
</tr>
<tr>
<td>2012</td>
<td>0.08987</td>
<td>0.06389</td>
<td>0.08125</td>
<td>0.08476</td>
<td>0.31977</td>
</tr>
<tr>
<td>2013</td>
<td>0.04944</td>
<td>0.08226</td>
<td>0.09158</td>
<td>0.07470</td>
<td>0.29799</td>
</tr>
<tr>
<td>2014</td>
<td>0.10274</td>
<td>0.09350</td>
<td>0.08468</td>
<td>0.05746</td>
<td>0.33838</td>
</tr>
<tr>
<td>2015</td>
<td>0.14832</td>
<td>0.10586</td>
<td>0.09458</td>
<td>0.12144</td>
<td>0.47020</td>
</tr>
<tr>
<td>2016</td>
<td>0.18139</td>
<td>0.11961</td>
<td>0.00689</td>
<td>0.07312</td>
<td>0.38101</td>
</tr>
<tr>
<td>2017</td>
<td>0.16300</td>
<td>0.13666</td>
<td>0.05682</td>
<td>0.06751</td>
<td>0.42400</td>
</tr>
<tr>
<td>2018</td>
<td>0.20196</td>
<td>0.16022</td>
<td>0.10212</td>
<td>0.16487</td>
<td>0.62917</td>
</tr>
</tbody>
</table>
Figure 1: Performance level of marine ecological civilization construction and comprehensive development index of each subsystem in Guangdong Province from 2009 to 2018

As can be seen from Table 2 and Figure 1, the performance level of Marine ecological civilization construction in Guangdong Province showed an overall upward trend from 2009 to 2018, increasing from 0.25338 in 2009 to 0.62917 in 2018, which increased by two times. This data reflects that Marine ecological civilization construction in Guangdong Province is developing in a sustainable and healthy direction. The overall development level of economic, social and scientific subsystems has been on the rise in the past decade. Guangdong Province attaches great importance to the development of economy, society and science and technology, but ignores the importance of the environment. Thus, it has made progress in economy, society, and science and technology. However, the development of the environment is gradually ignored, which leads to the imbalance of the overall development. Compared with the economic and technological subsystem, the overall development index of the social and environmental subsystem is slightly lower. Therefore, Guangdong province needs to pay attention to improving people's quality of life and strengthening environmental protection and management in the future construction process.

3.2 Coupling Coordination Analysis

According to the real data of Guangdong Province from 2009 to 2018 and Eqs. (8)-(10), the coupling coordination degree of the four subsystems of economy, society, environment and science and technology is calculated. The results are shown in Fig. 2 and Table 3.
The degree of coupling and coordination of marine civilization construction in Guangdong Province has increased in the past ten years. From 0.13770 in 2009 to 0.86639 in 2018. From 2011 to 2014, the coupling degree is relatively stable, from low coupling stage to high coupling stage. And the coupling degree is relatively stable from 2012 to 2015, which shows that in the past four years, there is a strong correlation between the economic, social, environmental and scientific subsystems. From 2015 to 2016, the direct economic loss index of marine disasters decreased from 2.877 billion yuan to 963 million yuan, and there was a gap between the speed of economic development and the other three systems. The correlation between the four subsystems weakened, and the coupling degree decreased from 0.92099 in 2015 to 0.63375 in 2011.

The degree of coupling coordination increased from 0.10878 in 2009 to 0.40613 in 2018, and the type of coupling coordination increased from basic imbalance to good coordination. Among them, 2009–2011 is the basic coordination stage, 2012 is the intermediate coordination stage, 2013 is the basic coordination stage, 2014–2015 is the intermediate coordination stage, 2016 is the basic coordination stage, 2017 is the intermediate coordination stage, and 2018 is the good coordination type.
Figure 2 shows that the coupling coordination degree of the performance evaluation system of marine ecological civilization construction in Guangdong Province is low. There are differences among the four subsystems of economic, social, environmental and scientific, and technological development. On the whole, the coupling coordination degree shows a slow upward trend, that is, the difference between subsystems in the construction process decreases, showing a general direction of development.

4 CONCLUSIONS

We select 14 indicators for evaluation and analysis from four aspects: economy, society, environment, science and technology. The conclusions are as follows.

From 2009 to 2018, the marine economy of Guangdong Province showed an increase, and the industrial structure continued to be optimized. The marine GDP showed linear growth, and the economy continued to develop. With regard to marine scientific research institutions, more experts are engaged in marine research, and the ability of marine scientific research is increasing. In the follow-up process, it is necessary to pay attention to environmental protection and restoration and strive to improve people's living standards and quality of life.

The overall performance level of marine ecological civilization construction in Guangdong Province shows an increase but the comprehensive development index of social environment is low. Therefore, Guangdong Province needs to pay attention to improving people's quality of life and strengthening environmental protection and governance in the process of marine ecological civilization construction in the future.

From 2009 to 2018, the coupling coordination degree of Guangdong marine ecological civilization showed a slight increase from 0.10878 in 2009 to 0.40613 in 2018 from the basic coordination stage to the good coordination stage. The coupling degree increased from 0.13770 in 2009 to 0.86639 in 2018. The increase from low level coupling stage to high level coupling stage indicates that the differences among the four subsystems of economic, social, environmental and scientific development are decreasing and developing in a right direction on the whole.

Acknowledgements

This research was supported by the Guangdong Philosophy and Social Science Planning Project (Research on the trading mechanism of coastline resources in Guangdong - Hong Kong - Macao Greater Bay Area from the perspective of marine ecological civilization construction, Project No. GD20XGL14)

This research was supported by the Humanities and Social Sciences Research Planning Fund of the Ministry of Education (Research on dynamic pricing and trading mechanism of coastline resources in Guangdong-Hong Kong-Macao Greater Bay Area under the background of marine ecological civilization construction, Project No. 21YJA630097)
REFERENCES


The Innovation and Application of Big Data Technology in Cross-Border E-Commerce Comprehensive Test Zone in China

Lixia Hu
hulixia77@sina.com
Beijing College of Finance and Commerce, Tongzhou, Beijing, China

Abstract: The digital transformation of the world economy is accelerating, a new round of scientific and technological revolution and industrial transformation is deepening, and the iterative upgrading and converged application of technology driven by e-commerce continues to deepen. The cross-border e-commerce comprehensive test zone in China is the “Gathering place” of system innovation, management innovation, service innovation and technology innovation, in particular, in the extensive application of big data and other information technologies to promote the reform and development of the comprehensive test zone to form a number of worthy reference and promote the application of the advanced experience, it includes building new statistical model, risk early warning system, intelligent logistics system, and global trade precision marketing, to promote the high-quality development of global digital trade.

Keywords: Cross-Border E-Commerce Comprehensive Test Zone, Big Data Technology, Statistical Index, Risk Early Warning, Intelligent Logistics, Precision Marketing.

1 INTRODUCTION

1.1 The development course of cross-border e-commerce comprehensive test zone in China

The cross-border e-commerce comprehensive test zone in China started in 2012, it has gone through the development stages of the pilot phase (2012-2014), the first batch (March 2015), the second batch (January 2016), the third batch (July 2018), the fourth batch (December 2019), the fifth batch (May 2020), the sixth batch (February 2022), and the seventh batch (November 2022). As of December 2022, it has built 165 cross-border e-commerce comprehensive test zones, covering 31 provinces, municipalities and autonomous regions nationwide, the quantity, development scale and regional distribution of the comprehensive test zone have all gone up to a new stage. With the number of approved cities as the standard, we can divide the cross-border e-commerce comprehensive test zone in China into three echelon. The first echelon is: more than 10 cities, the top four were Guangdong Province, Shandong Province, Jiangsu province and Zhejiang province. The second echelon, there were less than 10 cities and more than five cities, in terms of number from more to less, they include Jiangxi province, Fujian Province, Sichuan Province, Anhui province, Liaoning
province, Hunan Province, Hebei province, Henan province and Inner Mongolia Province. The third echelon, less than 5 cities, in terms of number from more to less, they include Hubei Province, Heilongjiang Province, Jilin Province, Xinjiang Province, Yunnan Province, Guangxi Zhuang Autonomous Region and so on, the detailed distribution is shown in Table 1.

**Table 1** The distribution table of “Three echelon” of cross-border e-commerce comprehensive test zone in China

<table>
<thead>
<tr>
<th>Area</th>
<th>Quantity</th>
<th>Echelon</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guangdong Province</td>
<td>21</td>
<td>the first echelon</td>
</tr>
<tr>
<td>Shandong Province</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Jiangsu province</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>Zhejiang province</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Jiangxi province</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Fujian Province</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Sichuan Province</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Anhui province</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Liaoning province</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Hubei Province</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Inner Mongolia autonomous region</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Hebei province</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Xinjiang autonomous region</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Yunnan Province</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Guangxi Zhuang autonomous region</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Guizhou Province</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Shanxi Province</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Shaanxi Province</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Hainan Province</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Gansu Province</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Qinghai Province</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Beijing</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Chongqing</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Shanghai</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Tianjin</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Ningxia Hui autonomous region</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Tibet</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>165</td>
<td>--</td>
</tr>
</tbody>
</table>
1.2 The present research situation of cross-border e-commerce comprehensive test zone in China

Using the keywords “Cross-border e-commerce comprehensive test zone” as the title, we can retrieve 248 research results from 2015 to 2022 on CNKI, this is shown in Figure 1.

In 2015, Hangzhou cross-border e-commerce comprehensive test zone as the first cross-border e-commerce comprehensive test zone in China was formally established, the same year there are related topics of research results. During the 8 years from 2015 to 2022, research results showed an overall upward trend, with the highest peak in 2020 and a slight decline in 2021 to 2022. The research topics mainly focus on: government policy, development strategy, Hangzhou (Guangdong, Henan, Shandong, etc.) city construction experience, tax policy, industry competitiveness evaluation.

Analysis of research outcomes 2015-2016: with the development of the first batch of cross-border e-commerce comprehensive test zone and the establishment of the second batch of 12 urban comprehensive test zone, the research results of these two years are mainly carried out around the mode, experience and advantages and disadvantages of the construction of the existing comprehensive test areas, representative achievements include standardization promoting the development of the Hangzhou cross-border e-commerce comprehensive test zone (Zhang Xin et al., 2015), actively linking the “Belt and Road” strategy, and accelerating the construction of the China (Dalian) cross-border e-commerce comprehensive test zone (Yang Xiaomeng et al., 2016).

Analysis of research outcomes 2017-2019: the results of this stage will continue to revolve around the four groups of cross-border e-commerce comprehensive test zone that have been established, and continue to summarize their experience in development and construction, in addition, the depth and breadth of research has been new changes. SWOT analysis, new trade model, big data, evaluation of industrial competitiveness, some new research methods, models and technologies are gradually combined with the development of cross-border e-commerce comprehensive test zone. Representative research results include: development strategy research of Zhengzhou cross-border e-commerce comprehensive test zone based on SWOT analysis (Zhang Juntao, 2018), cross-border e-commerce and the coordinated development of the Economic system-an empirical study based on 35 cross-border e-commerce comprehensive test zone (Zhang Xiaodong, 2019).

Analysis of research outcomes 2020-2022: the systematic results in development policy analysis of cross-border e-commerce comprehensive test zone, high-quality research results have also been achieved in the qualitative and quantitative analysis of sample data from the
cross-border e-commerce comprehensive test zone, this paper sums up the problems of the mode, performance, supply chain development potential and other important fields of cross-border e-commerce comprehensive test zone in China, and forms a conclusion of great reference and reference value. The representative research results include: the research on evolution dynamics and innovation realization mechanism of cross-border e-commerce comprehensive pilot area (Xiao Liang et al., 2020), mechanisms and effects of quality change in manufacturing enabled by digital trade-quasi-natural experiments from the cross-border e-commerce comprehensive test zone (Yuan Qigang et al., 2022).

1.3 The significance of this study
The digital transformation of the world economy is accelerating, a new round of scientific and technological revolution and industrial transformation is deepening, and the iterative upgrading and converged application of technology driven by e-commerce continues to deepen. The cross-border e-commerce comprehensive test zone in China is a “Gathering place” for institutional innovation, Management Innovation, service innovation and technological innovation, through a series of innovations to promote the cross-border e-commerce industry liberalization, facilitation, standardization of development, thus promoting the higher-quality economic development in China. The cross-border e-commerce comprehensive test zone in China, in the extensive application of big data and other information technology, to promote the reform and development of the comprehensive test zone to form some valuable experience.

2 MAKE FULL USE OF BIG DATA TO ESTABLISH NEW MODELS AND SYSTEMS FOR CROSS-BORDER E-COMMERCE STATISTICS

The cross-border e-commerce comprehensive test zone in China actively utilizes information technologies such as big data, artificial intelligence, blockchain, cloud computing, etc. They use new technologies to promote service innovation, Control risks and regulate the smooth operation and development of the economy in the comprehensive test zone. Hangzhou, Shenzhen, Zhengzhou, Ningbo, Yiwu, Qingdao and other cross-border e-commerce comprehensive test zones have made full use of big data and other technologies to establish new models and systems for cross-border e-commerce statistics.

2.1 Set up and publish cross-border e-commerce index
Using new technologies such as big data and cloud computing to analyze and process huge amounts of data on various platforms, such as commodity trading, logistics, customs clearance, financial payments, intellectual property rights, etc., it will gradually establish a comprehensive index system that reflects the operation of cross-border e-commerce at multiple levels and dimensions, and regularly release the “Cross-border e-commerce index” to the whole society, to guide and monitor the economic development and smooth operation of the comprehensive cross-border e-commerce comprehensive test zone.
2.2 Establishing a new model for cross-border e-commerce statistics

To establish a cross-border e-commerce big data service center to realize the exchange and aggregation of cross-border e-commerce data, and to explore the establishment of a new model for statistics and management based on declaration lists, platform data, etc. Establish a “Cross-border e-commerce data monitoring system” to provide decision-making advisory services for government regulation and business operations.

2.3 Establish statistical standards for cross-border e-commerce

To explore the establishment of standard formats such as information of transaction entities, electronic contracts and electronic orders, as well as simplified statistical classification standards for cross-border e-commerce imports and exports, and to explore the establishment of a multi-party cross-border e-commerce statistics system, to improve cross-border e-commerce statistical methods, and provide experience for national cross-border e-commerce statistical system and mechanism construction.

3 APPLICATION OF BIG DATA TECHNOLOGY, THE ESTABLISHMENT OF RISK EARLY WARNING SYSTEM

3.1 Establish risk early warning system based on big data

Ningbo cross-border e-commerce comprehensive test zone supports third-party credit service providers to provide credit evaluation services to governments and enterprises through big data technologies; and establishes an early warning system for classified and graded risks based on big data analysis, to realize real-time monitoring, identification and evaluation of market risks, transaction risks, financial risks, technical risks and public security risks in cross-border e-commerce.

3.2 Use big data and other technical means to innovate enterprise credit rating methods

Nanjing cross-border e-commerce comprehensive test zone proposes to build a big data trade finance platform, to use big data and other technical means to innovate enterprise credit rating methods, to provide non-face-to-face approval of trade finance products and services, to provide convenience for enterprises to carry out various types of financing business.

3.3 Use big data to achieve "point-to-point” matching and control enterprise financing needs and risks

Qingdao cross-border e-commerce comprehensive test zone fully mining cross-border e-commerce generated by internet trade big data, through data information dynamic and independent “Point-to-point” matching and hedging, to realize the dynamic control and effective control of enterprise financing demand and risk management. The wide application of new technologies such as big data and artificial intelligence has greatly improved the advanced service, scientific management and sensitivity of risk early warning.
4 APPLICATION OF BIG DATA AND OTHER TECHNOLOGIES, THE ESTABLISHMENT OF INTELLIGENT LOGISTICS SYSTEM

4.1 Rely on big data to promote the transformation of foreign trade from logistics oriented to information oriented

Qingdao cross-border e-commerce comprehensive test zone relies on technical services such as big data, internet of things and cloud computing to promote the transformation of foreign trade from logistics-oriented to information flow-oriented, and to provide cross-border e-commerce information services, to guide the integration of settlement, transportation and other related services, and promote the facilitation of cross-border e-commerce trade in goods.

4.2 Use the internet of things and big data technology to build an interconnected intelligent logistics information system

Hangzhou, Zhengzhou and Qingdao cross-border e-commerce comprehensive test zone make full use of new technologies such as the internet of things and big data to build interconnected intelligent logistics information systems, to standardize and standardize the operation process of cross-border e-commerce logistics by connecting and operating a smooth logistics warehousing network system, a high-quality and efficient logistics operation service system, etc., to integrate transport resources efficiently and seamlessly link up transport organizations to form a rational, full-featured, efficient and high-quality cross-border logistics distribution and operation service system.

5 APPLICATION OF BIG DATA TECHNOLOGY, INNOVATION OF INSPECTION AND QUARANTINE PROCESS

Guangzhou cross-border e-commerce comprehensive test zone applies big data technology to inspection and quarantine processes, relying on “Smart inspection ports” for risk assessment, classification management, Integrity Management and overseas information comparison, to realize “Electronic distribution and control”, explore the application of global product quality and standard information database for inspection and quarantine big data automatic judgment, automatic check and release. It will explore ways to control the quality of imported goods in cross-border e-commerce and encourage domestic enterprises to export famous-brand goods through market procurement. To formulate the administrative measures for the procedures of inspection and quarantine declaration and release in the cross-border e-commerce comprehensive test zone in Guangzhou, the supervisory model of “Pre-filing, pre-supervision, post-follow-up, quality control” shall be implemented for exit, and “Pre-filing and filing for entry, centralized inspection and quarantine in entry areas, verification and verification in batches in exit areas, and tracing of quality and safety” shall be implemented for entry. Full big data technology, product quality traceability management, to ensure the quality of import and export products, good product quality this pass.
6 APPLICATION OF BIG DATA TECHNOLOGY TO ACHIEVE GLOBAL TRADE PRECISION MARKETING

6.1 The model of “Big Data + e-commerce + foreign trade”

Zhengzhou cross-border e-commerce comprehensive test zone explores the development model of "Big Data + e-commerce + foreign trade" to gradually build Henan foreign trade big data center and provide customized services for enterprises, to help companies in the global market accurate marketing and rapid increase in overseas market share.

6.2 The mode of transformation from mass manufacturing to mass customization

Qingdao cross-border e-commerce comprehensive test zone uses big data technology to help foreign trade enterprises convert customers' individual needs into valuable orders in accordance with market changes, we will build a flat cross-border trading system featuring on-demand design, on-demand manufacturing and on-demand distribution, and realize the transformation of large-scale manufacturing to mass customization and the optimization of product imports and exports.

6.3 The mode of global trade precision marketing big data platform

Nanjing cross-border e-commerce comprehensive test zone in the construction of an international network marketing system, scientific planning, outstanding results. Its main measures include speeding up the distribution of overseas marketing network, implementing the action plan of "Breaking Zero", promoting accurate marketing in international market, promoting the upgrading of trade fairs, and implementing the strategy of winning by quality, for the development of the comprehensive test area of the international market to make a positive contribution. In particular, it will speed up the introduction and cultivation of third-party overseas marketing services to help enterprises develop overseas independent stations; To guide and encourage enterprises to carry out international e-commerce-related products and enterprise certification, to establish international product standards, to enhance the voice of international competition, and so on, it shows the foresight and precision of the government in the aspect of careful planning and scientific help.

7 CONCLUSIONS

The Fourteenth Five Year E-commerce Development Plan (October 2021) jointly issued by the Ministry of Commerce, the Central Cyberspace Office and the National Development and Reform Commission clearly pointed out that "continue to promote the construction of cross-border e-commerce comprehensive test areas and explore the innovation of the whole process of cross-border e-commerce transactions." In October 2021, Premier Li Keqiang said in his speech at the opening ceremony of the 130th China Import and Export Commodities Fair and the Pearl River International Trade Forum: "Rely on industrial and market advantages, deepen foreign economic and trade cooperation, accelerate the innovative development of foreign trade, strengthen international cooperation in digital trade, and create a number of digital pilot areas for global trade."
From the above analysis, we can see that the development of digital technologies such as big data, artificial intelligence and cloud computing, as well as the extensive application of innovation in cross-border e-commerce comprehensive test zones, not only promotes the optimization and upgrading of cross-border e-commerce supply chain services in China, but also accurately captures the needs of international market customers, realizes accurate marketing and effective marketing, and more effectively realizes digital supervision, to ensure the healthy development of cross-border e-commerce. In the wave of global digital development, cross-border e-commerce will gradually change into global digital trade with the accumulation of continuous quantitative changes, and its digital characteristics become more and more prominent, the issue of big data innovation and application in this industry is also becoming more important. It is hoped that the 165 cross-border e-commerce comprehensive test zones in China will have more innovation and exploration on the basis of the previous development experience, and form the "China Experience", which will contribute to the development of global digital trade.

**Fund Project:** This paper is one of the phased research results of the Research on Development Experience, Mode and Trend of China's Cross-border E-commerce Comprehensive Test Zone funded by R&D Program of Beijing Municipal Education Commission. (project No.: SM202051638001)

**The Author:** Hu Lixia, female, July 1977, Han nationality, Chicheng, Hebei province, Ph. D. , Associate Professor, Research Direction: market research and statistics, cross-border e-commerce, etc.

**REFERENCES**

[3] Implementation plan for the construction of 10 cross-border e-commerce comprehensive test zones in Hangzhou, Ningbo, Qingdao, Shanghai, Guangzhou, Shenzhen, Zhengzhou, Xiamen, Nanjing and Yiwu.


Application of Job Analysis in Recruitment Based on Big Data Technology: Take F Company as an Example

Li Hang¹, Huam Hon Tat²
lihang.1989@163.com¹, huam@putrabs.edu.my²
Infrastructure University Kuala Lumpur, Malaysia¹
Putra Business School, Malaysia²

Abstract: Job analysis provides essential support for corporate strategic planning and is the basis of human resource management. It can help organizations comprehensively analyze the work that needs to be performed. In this paper, the work method of job analysis is used to analyze the structure of the key work elements of the recruiting position. This meets the requirements of the company's general manager and the employment demand department, hoping to determine the key work elements of each position to improve the suitability of newly hired employees with the company.

Keywords: Job Analysis, Recruitment, Big Data, Human Resource Management.

1 INTRODUCTION

With the advent of the era of digital innovation, big data technology has received the most attention in various fields and has been applied in practical work. Its functions of information integration, mining, and analysis provide organizations with accurate and scientific information. Support (Xiao et al., 2022). Big data technology has also been well-practised in human resource management in enterprises. As the most important part of human resources, recruitment is one of the key fields where big data technology can be used (Zhang et al., 2020). Company F is a high-tech enterprise in China, mainly engaged in the R&D and manufacturing of marine engineering equipment. Although the company provided employees with market-competitive salaries and benefits, the performance of the recruited new recruits during the trial period did not satisfy the managers. In view of this situation, the general manager analyzed the various situations that exist at the moment. In order to verify the real cause of the problem, he discussed and communicated with the human resources department and the employment demand department, and asked the person in charge of the human resources department to provide an authentic analysis report, explaining the specific situation of the company's recruitment. The goal is to study the key elements in all aspects of recruitment, so as to improve the accuracy of recruitment and meet the actual needs of enterprises.
2 LITERATURE REVIEW

2.1 Job analysis

Job analysis allows the company to clearly understand the job content, job responsibilities, and qualifications of employees in their positions, and allows managers to clearly see the workflow between various departments and positions (Andhi et al., 2019). In recent years, many companies have begun to pay attention to the importance of recruitment, how to recruit the right people, and find a way to quickly recognize the company's corporate culture with high-fit level employees, so as to achieve a win-win situation for the company and employees (Nasution et al., 2019).

Job analysis is a method of comprehensively analyzing a certain job or task, and will determine its key elements. During the process, different research methods can be used for the job analysis (Nasution et al., 2020), such as the questionnaire method, interview law etc. Each method has advantages and disadvantages. Different research methods are used for different situations so that key information can be found more accurately. In complex cases, different research methods may be used for analysis work. The determination of research methods depends on the researcher's work experience and professional knowledge has higher requirements.

Job analysis has made an important contribution to recruitment work, and has attracted the attention of the company's management. Applying scientific and technical means to analyze the recruited positions, so as to determine the key work elements, targeted work can improve efficiency and reduce manpower Resource input (Lohman et al., 2020).

2.2 Big data

In the summary of the characteristics of big data, many experts believe that big data has the characteristics of Value, Volume, Velocity, Variety, etc., combined with the definition of big data by domain experts, we can conclude that we need to collect a large amount of data, through a certain logic Computational methods, analysis and classification of the integrated data (Yin et al., 2020).

The difference between big data technology and previous data is that the amount of data information is larger, the analysis is more precise, and the integration is more comprehensive (Xiao et al., 2022). Big data analysis has broken through the limitations of previous data analysis. With the rapid development of digital information technology, the amount of global data is growing at an unprecedented rate. People can refer to more data when dealing with work. Not only can they learn from previous information, but they can also obtain the latest and new information and data at the first time, conduct analysis more objectively, and form pertinent evaluations and solutions (Zhang et al., 2019).

Under the application of big data technology, the human resource management of enterprises collects information through the network information base, and provides decision-making for the formulation of strategies after analysis. The data-driven approach makes human resources work faster and saves time, and gradually drives the improvement of the work efficiency of other departments of the enterprise, providing useful help for the better and faster development of the enterprise (Zhang et al., 2020).
3 RESEARCH METHODOLOGY

This study adopts the technology of determining key work elements, conducts qualitative and semi-structured interviews with the heads of employment demand departments, uses questionnaires to collect information on employees in the same position, and reflects the work background based on big data. The application of job analysis in recruitment, the steps of job analysis are as follows: 1. Work breakdown structure (analyze data, organize information on key elements of work), 2. Key work elements are identified based on the findings. This article describes only one position—administrative specialist as an example—and identifies key elements of its job, since the analysis of other jobs is the same.

3.1 Determine the key elements of the job

First, determine the main work elements, such as asset procurement, conference reception, canteen management, vehicle management, etc.

Secondly, the key elements of each job of the post are recorded in a rough form, as shown in Figure 1:

![Figure 1: Key job elements of an administrative commissioner](image)

3.2 Sorting of key work elements

Rank the key elements of work, depending on the following factors (based on previous experience):

- Priority (urgent/normal/not urgent)
- Importance (important/normal/not important)
- Nature (difficult/normal/easy)

3.3 Key elements of qualifications

The first choice, determine the requirements of the position for applicants, such as work experience, academic background, professional skills, and attitudes.

Secondly, each qualification element of the post is recorded in a rough form, as shown in Figure 2

![Figure 2](image)
3.4 A collection of key elements for the recruitment of administrative specialists

According to the opinions and answers of the person in charge of the employment demand department and the employees in the same position, the results of job analysis are obtained, as shown in Figure 3:

3.5 Application of job analysis results in big data technology

Through job analysis, it is found that the key work elements of the post are the application layer, which is entered into the recruitment platform; the platform system is the functional layer, and the data management module, logic model module, and data analysis module of the system start to be based on the requirements of the application layer. The search is started in the database, and the processed data is presented to the presentation layer. The display layer is
the user interface of the recruitment platform. According to the data model, analyzed and searched by the system, there is a logical and scientific basis to provide professional data support for the company's recruitment work. The application logic diagram is as follows:

![Figure 4: The use of big data technology in recruitment](image)

### 3.5.1 User level - identification of key job elements

Company F uses job analysis to determine the key job factors for recruitment positions according to the different job settings and job qualifications. Company F draws up qualitative semi-structured interviews with the head of the employment needs department and questionnaires with employees in the same position, organizes the results of the analysis according to the set forms and publishes the recruitment information with the help of the recruitment website so that the recruitment needs spread quickly and effectively. At the same time the recruitment website will also be based on the analysis of the recruitment intentions of the job seekers and the positions submitted, using the algorithm of big data, the recruitment will be recommended to job seekers who are interested and meet the recruitment requirements, the establishment of a special talent pool, the system automatically communicates with the talent through private messages to reduce human costs, the logic diagram is as follows:

![Figure 5: Application level logical relationships](image)
3.5.2 Functional level - analysis and screening of CVs

After the information is entered into the recruitment platform, according to the screening of key job elements in the database, after big data analysis, the results are derived, partly pointing to the user and partly pointing to the enterprise, the enterprise receives information that meets the conditions of the key job elements provided by the enterprise, while the enterprise takes the initiative to search, big data will recommend the number of resumes with similar key elements to the enterprise, reducing the imprecision of the results of job analysis due to The number of CVs that are missed due to inaccurate job analysis results is reduced. In general, after two or more key job factors are met, the active recommendation of CVs is then analysed and similar CVs are recommended. The functional level logic diagram is as follows:

![Functional level logical relationships](image)

Figure 6: Functional level logical relationships

3.5.3 Presentation Layer - Selection of Candidates

For the job seekers recommended and screened by the system, the staff of the human resources department need to further accurately identify and judge their matching degree with the recruiting position. For users actively recommended by the system, the matching degree of their big data analysis can be used as a reference.

4 BIG DATA ANALYTICS IN RECRUITMENT

Big data analysis technology not only collects, screens, analyzes and integrates data from the recruitment platform, but also needs to be analyzed according to the degree of informatization, the company's development strategy, recruitment needs, and the professional ability of recruiters. Only by setting plans can the effective combination of big data technology and enterprise recruitment be realized.

4.1 Selection of recruitment platform

The effect of big data analysis technology in recruitment is closely combined with the needs of enterprises. The selection of enterprise recruitment platforms should be classified into four aspects according to the industry, region, company nature and job requirements. The effectiveness of recruitment depends on the reliability and accuracy of the data source of the recruitment platform. The database of each recruitment website has its own characteristics and advantages. Some are good at high-end management posts, some are good at professional
technical posts, and some are more suitable for recruiting workshop workers. Choosing a recruitment website with a high degree of fit with the needs of the enterprise can ensure more data and accurate data analysis. A suitable recruitment website cannot guarantee the accuracy of big data analysis, and it can recruit more talents to fill the database, which will make the company's recruitment information spread and disseminate more quickly, and select qualified job seekers to the company, so that the recruitment done more quickly and efficiently (Dong et al., 2020).

4.2 The relationship between big data analysis and job analysis results

Big data analysis technology requires enterprises to provide information scope and keywords, and the result of job analysis is the main analysis tool for enterprises to provide precise screening conditions. The big data analysis algorithm of the recruitment platform will calculate the results according to the screening requirements, so the professional ability of the company's human resources practitioners also has an important impact on the results of big data analysis.

5 THE IMPACT OF BIG DATA ANALYTICS ON RECRUITMENT RESULTS

Company F's recruitment methods are mainly divided into on-site recruitment and online recruitment. On-site recruitment is mainly for campus recruitment and offline job fairs held by the government, which are staged special recruitment. Considering labor cost and time cost, Company F prefers to use online recruitment.

5.1 Advantages and disadvantages of on-site recruitment

5.1.1 Advantages

a) You can directly communicate with the interviewers face to face, as a preliminary test, control the number of job seekers, and have a preliminary impression of the quality.

b) Campus recruitment is generally free, which can save costs.

5.1.2 Disadvantages

a) The labor cost is high, and a large number of employees are usually required to be on site.

b) The number of people participating in the on-site recruitment is often limited to the propaganda of the sponsor, which has a great influence and restriction on the quantity and quality of job seekers.

c) It is impossible to analyze resumes with analysis tools, and the accuracy is poor.

d) are often influenced by first impressions.
5.2 Advantages and disadvantages of online recruitment

5.2.1 Advantages

a) The coverage is wide, allowing more job seekers to see the recruitment needs of the company.

b) Strong timeliness, the recruitment method is interactive, not limited by time and region, it can transmit information more quickly and conveniently, and the update speed is also very fast.

c) More precise, the big data analysis function will integrate and match the needs of the enterprise and job seekers through the data in the database.

5.2.2 Disadvantages

a) The authenticity of information is low, and the reliability of verification such as academic qualifications is low.

b) The position limitations need to be expanded. For manual labor positions, such as production workers, drivers, chefs, etc., the recruitment effect is not ideal because they are not skilled in using the Internet.

6 DISCUSSION

In this paper, the work method of job analysis is used to analyze the structure of the key work elements of the recruiting position. The online recruitment channel of Company F is the Zhaopin platform, and the results of the key work elements are entered into the Zhaopin platform as screening conditions. More accurate screening of resumes and resumes searched online. This meets the requirements of the company's general manager and the employment demand department, hoping to determine the key work elements of each position to improve the suitability of newly hired employees with the company.

In the recruitment process, the ZhiLian recruiting platform uses big data technology combined with the key work elements of job analysis and summary, makes full use of its powerful resume search engine and a large amount of resume data, draws the talent portrait of the company's recruitment position, and conducts job analysis for the position. Design a competency model and screen resumes based on big data algorithms to achieve a scientific matching effect.

In the analysis, determining the key work elements cannot cover all the work content of the post, and some temporary work elements are ignored because they do not account for a high proportion of the workload. They found that job analysis to determine the key elements of the job, with the help of big data technology, can help employees to better complete the company's recruitment tasks in a professional, accurate, and fast manner (Shao al., 2020). These are typical recruiting job-related traits that are crucial in any job and contribute to good results for employees in a business. Therefore, these characteristics have been emphasized in the study of this paper.
Job analysis is an important tool in human resource management. With the help of big data technology, it is conducive to improving the accuracy and effectiveness of recruitment work. It can help companies save recruitment costs, improve work efficiency, and enable employees to quickly integrate into the enterprise, to achieve a win-win situation between the enterprise and employees. This article takes a company whose recruitment results are not satisfactory as an example and discusses how a systematic approach can recruit more suitable employees through merit and demerit job analysis in the context of big data technology. The analysis done in this paper can serve as a sample of those organizations that need some investment in this regard.

REFERENCES


Research on the Evaluation of the Transformation Capability of Scientific and Technological Achievements of Chinese Universities Based on Entropy Weight Method and Cluster Analysis

Jing YIN¹, Guanghui ZHU², Mingle ZHOU²,*

*zhouml@qlu.edu.cn

Department of Economics and Management, Qilu University of Technology (Shandong Academy of Sciences), Jinan 250353, China¹
Shandong Computer Science Center (National Supercomputer Center in Jinan), Qilu University of Technology (Shandong Academy of Sciences), Jinan 250014, China²

Abstract: The transformation of scientific and technological achievements is crucial for the combination of science and technology and economy. Universities are important subjects in the transformation of scientific and technological achievements, which are crucial to promoting economic development and social progress. From the perspective of the transformation process of scientific and technological achievements, this study constructs an evaluation index system to assess the level of science and technology achievement conversion in Chinese universities. The research data was collected from the "Compilation of Science and Technology Statistics for Higher Education Institutions in 2021" published by the Department of Science and Technology, Ministry of Education, China. The entropy weight method and cluster analysis were used to evaluate the transformation capability of scientific and technological achievements of universities in 31 provinces and cities (except for Hong Kong, Macao and Taiwan) in China. Research indicates that there are large differences in the transformation capability of scientific and technological achievements among Chinese universities in different provinces and cities. In addition, universities in some provinces and cities have uneven development levels in various stages of achievement transformation process. The theoretical and practical implications of this study are presented at the end.

Keywords: Universities, Transformation Of Scientific And Technological Achievements, Transformation Capacity Of Achievements, Entropy Weight Method, Cluster Analysis.

1 INTRODUCTION

The transformation of scientific and technological achievements is a key link of scientific and technological innovation. As an important subject to promote and implement the activities of transformation of scientific and technological achievements, universities' ability to enhance the transformation of scientific and technological achievements plays an important role in promoting the sustainable growth of national economy. Thus, it is essential to systematically evaluate the capability of universities to transform scientific and technological achievements.
In addition, the evaluation can identify the problems in the process of transformation in higher education institutions. This is significant for exploring effective ways for universities to improve the transformation capability of scientific and technological achievements.

The transformation of scientific and technological achievements refers to the process of transforming scientific and technological achievements from laboratories and research centers into useful products and services to meet market demand and achieve economic and social benefits. The term "transformation of scientific and technological achievements" is a unique term in China's science and technology management. Internationally, scholars mostly use the concepts of "technology transfer", "technological innovation" and so on [15]. The scope of existing research on technology transfer is very extensive. However, as a whole, the existing literature mainly focuses on the patterns of technology transfer [2, 4-5], the factors that constrain technology transfer [4-5], and the evaluation of technology transfer efficiency [1, 6]. Few studies have been conducted in the literature from the perspective of technology transfer capability. In order to fill the gaps of existing studies, this research constructs an assessment indicator system, aiming to evaluate the technology achievement transformation capability of Chinese universities. This study attempts to find out the questions existing in the conversion of scientific and technological achievements of universities in various regions, and accordingly gives directions and suggestions for improvement to help universities improve their conversion capability.

2 LITERATURE REVIEW

The existing literature has studied the transformation of scientific and technological achievements mainly around three aspects. First, the existing literature has studied the modes and ways of transformation of scientific and technological achievements. Moortel and Crispeels [9] compare the patterns and cooperation paths of technology transfer conducted by Chinese and Western economies on the basis of constructing a theoretical framework for international intertechnical transfer. They found that technology transfer between international is mainly carried out in two ways: formal and informal ways. Baglieri et al. [2] conducted a study by collecting data from 60 universities in the United States and found that the main modes of technology transfer conducted by universities were university-enterprise partnerships, entrepreneurial teaching, patent transfer and business incubation.

Second, the existing literature has empirically studied the efficiency of technology transfer in research subjects. Rory et al. [14] empirically analyzed the relationship between knowledge spillover performance and university-enterprise alliances and technology transfer in U.S. universities around talent, resources, institutions, and finance. Danquah [7] used stochastic frontier analysis to study the efficiency of technology transfer in sub-Saharan African countries from 1970 to 2010. They found that human capital, research and development, and trade openness positively affect the efficiency of achievements transfer.

Third, the existing literature has explored the significant factors that influence the transformation of scientific and technological achievements. Das [8] studied the efficiency of technology transfer in 15 regions by constructing a computable general equilibrium model. Human capital, level of economic development and regional management capacity were found to be the main factors affecting technology transfer. Madanmohan et al. [12] found that R&D
investment and availability of skilled personnel, transfer pathways, government involvement and learning culture of firms influenced technology transfer capability.

By combing existing literature, it is found that the existing studies mainly focused on the transfer efficiency in the evaluation of the achievements transfer. Relatively little literature has studied the transformation capacity of scientific and technological achievements. Therefore, to make up for the deficiencies of existing literature, this research conducted a comprehensive evaluation of the transformation capability of scientific and technological achievements of universities in each region of China.

3 CONSTRUCTION OF INDEX SYSTEM

In the university, the activity of science and technology transformation is a complex process of many elements input, output, transfer and application. It can be mainly divided into three stages: project development, result output, and transfer application \(^{[13]}\). The evaluation of scientific and technological achievement transformation activities should cover the whole process of achievement transformation. In the evaluation, the index system should be constructed according to the characteristics and rules of universities. The constructed index system should reflect both the existing strength and development potential of scientific and technological innovation of universities, as well as the ability to transform scientific and technological achievements. Therefore, this study constructs the evaluation index system based on whole process of transformation of scientific and technological achievements. The level of scientific and technological research and development, the level of scientific and technological output and the level of transformation of scientific and technological achievements are primary indicators of the index system. They can reflect the strength of science and technology achievement transformation of universities comprehensively.

A school's research projects, research personnel, and financial investment are important indicators of its research strength and social status. This study selected science and technology manpower, science and technology funding, science and technology projects, and international science and technology exchanges as secondary indicators from the statistical compilation. Specifically, it includes 11 tertiary indicators such as the number of teaching and research personnel, the number of research and development personnel, internal and external expenditure funds, international academic conferences and so on. These indicators reflect the level of scientific and technological innovation in the process of transformation of scientific and technological achievements of universities, and reflect the competitiveness and R&D level of universities.

Scientific and technological achievements are the fruits of creative labor formed by researchers after scientific and technological research and development. The number of scientific and technological achievements determines the level of scientific research output of universities. Therefore, this study selected the scientific and technological achievements in the statistical data compilation as the secondary indicators of the second stage evaluation. The main forms of scientific and technological achievements are scientific papers, scientific works, invention patents and so on. Nine indicators, such as published scientific and technical works, academic papers and the number of patents granted, were selected as the tertiary indicators of
the evaluation system, so as to reflect the level of scientific and technical output of universities.

The transfer capability reflects the activity process of science and technology output and input among countries, regions, industries and within the system of science and technology itself. Based on the statistical data in compilation and the principle of accessibility of indicators, present study mainly measures the level of transformation of scientific and technological achievements through the revenue obtained from transfer of results. Specifically, it includes 3 tertiary indicators: the actual income in the year of patent sale, the amount of technology transfer contract and the actual income in the year of technology transfer. The evaluation index system constructed in this research is presented in Table 1.

4 METHODOLOGY

4.1 Data Collection and Instrument

The data used in this research was collected from the "Compilation of Science and Technology Statistics of Higher Education Institutions in 2021" published by the Department of Science and Technology, Ministry of Education, China. This compilation details the overall status of science and technology activities of 2078 higher education institutions and their affiliated hospitals in China in 2020. The content covers the situation of science and technology manpower, science and technology funding, science and technology projects, science and technology output and so on. In addition, this compilation does not include data on the science and technology activities of higher education institutions in Taiwan, Hong Kong and Macao of China. Therefore, statistics from 31 provinces and cities, excluding Hong Kong, Macao and Taiwan, were

<table>
<thead>
<tr>
<th>Primary indicators</th>
<th>Secondary indicators</th>
<th>Tertiary indicators</th>
<th>Variable name</th>
<th>Relevance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level of scientific and technological research and development</td>
<td>Science and technology manpower</td>
<td>Number of teaching and research staff (Person)</td>
<td>X1</td>
<td>Positive</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Number of R&amp;D staff (Person)</td>
<td>X2</td>
<td>Positive</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Number of R&amp;D results application and science and technology service personnel (Person)</td>
<td>X3</td>
<td>Positive</td>
</tr>
<tr>
<td>Science and technology funding</td>
<td>Internal and external expenditure funding (1000 yuan)</td>
<td>X4</td>
<td>Positive</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Research and development project current year expenditure funds (1000 yuan)</td>
<td>X5</td>
<td>Positive</td>
<td></td>
</tr>
<tr>
<td></td>
<td>R&amp;D results application current year expenditure funds (1000 yuan)</td>
<td>X6</td>
<td>Positive</td>
<td></td>
</tr>
</tbody>
</table>
4.2 Entropy Weight Method

In information theory, information entropy is a measure of the degree of uncertainty and disorder \[^{10}\]. The principle of entropy method of assigning weights is to determine the indicator weights based on the amount of information contained in the observations of each indicator. The more information, the less uncertainty. The smaller the entropy value of the index, the greater its influence on the comprehensive evaluation \[^{16}\]. In this research, the entropy weight method is used to assign the indicator weights to reduce the influence of subjective judgments, which makes the research findings out more objective and precise.

4.3 Cluster Analysis Method

Clustering is a classification method that groups data into clusters with similar characteristics. Data in clusters have a higher similarity, while data between clusters have a lower similarity...
Systematic clustering is one of the most commonly used clustering methods. Systematic clustering method is to split or aggregate data according to certain data connection rules, certain hierarchical structure, and finally form a hierarchical sequence of clustering solutions. The basic idea of this method is to cluster variables that are close to each other into classes first according to their distance, and variables that are farther away into classes later. This is done sequentially until each variable is grouped into the appropriate class \(^{(11)}\). In multivariate statistical analysis, the purpose of dimensionality reduction can also be achieved by clustering analysis.

5 ANALYSIS

5.1 Entropy Weight Method Analysis

The method used in this study is the entropy weight method, and the specific calculation procedure is shown below \(^{(16)}\).

First, the indicators are selected. Suppose there are \(m\) evaluation objects and \(n\) indicators, then \(x_{ij}\) represents the evaluation value of the \(j\)th indicator of the \(i\)th evaluation object \((i=1,2,\ldots,m; j=1,2,\ldots,n)\).

Second, the data standardization process. In this research, the extreme value method was applied to dimensionless each tertiary index to eliminate the difference in magnitude between different units. Each evaluation indicator can usually be divided into positive and negative indicators due to their different characteristics. The higher the value of the positive indicator, the better the evaluation. On the contrary, the smaller the value of the negative indicator, the better the evaluation. Therefore, different metrics require different normalization algorithms to process data. For positive indicators, it is necessary to use equation (1). For the negative indicator, it needs to be processed by equation (2). Table 1 shows that \(X_1\) to \(X_{23}\) are all positive indicators. Thus, the data were normalized in this study using equation (1).

\[
X_{ij} = \frac{x_{ij} - \min \{x_j\}}{\max \{x_j\} - \min \{x_j\}} \quad (1)
\]

\[
X_{ij} = \frac{\max \{x_j\} - x_{ij}}{\max \{x_j\} - \min \{x_j\}} \quad (2)
\]

Third, the weight of the \(i\)th evaluation object in the \(j\)th indicator is determined using equation (3).

\[
Y_{ij} = \frac{x_{ij}}{\sum_{i=1}^{m} x_{ij}} \quad (3)
\]

Fourth, the information entropy of the \(j\)th indicator is calculated using Equation (4).

\[
e_j = -\frac{1}{\ln m} \sum_{i=1}^{m} \left( Y_{ij} \ln Y_{ij} \right) \quad (4)
\]
Fifth, the information utility value of the $j$th indicator is calculated using equation (5).

$$d_j = 1 - e_j$$  \hspace{1cm} (5)

Sixth, the weight of the $j$th indicator is calculated using equation (6).

$$W_j = \frac{d_j}{\sum_{j=1}^{n} d_j}$$ \hspace{1cm} (6)

**Figure 1:** Dendrogram.

**Table 2:** Entropy value and entropy weight of each index.

<table>
<thead>
<tr>
<th>Variable name</th>
<th>Indicator name</th>
<th>Entropy value</th>
<th>Entropy weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td>Number of teaching and research staff</td>
<td>0.9342</td>
<td>0.0179</td>
</tr>
<tr>
<td>X2</td>
<td>Number of R&amp;D staff</td>
<td>0.9167</td>
<td>0.0226</td>
</tr>
<tr>
<td>X3</td>
<td>Number of R&amp;D results application and science and technology service personnel</td>
<td>0.8697</td>
<td>0.0354</td>
</tr>
<tr>
<td>X4</td>
<td>Internal and external expenditure funding</td>
<td>0.8612</td>
<td>0.0377</td>
</tr>
<tr>
<td>X5</td>
<td>R&amp;D project current year expenditure funds</td>
<td>0.8334</td>
<td>0.0452</td>
</tr>
<tr>
<td>X6</td>
<td>R&amp;D results application current year expenditure funds</td>
<td>0.7600</td>
<td>0.0652</td>
</tr>
</tbody>
</table>
Seventh, the composite score of the ith evaluation object is calculated using equation (7).

\[ S_i = \sum_{j=1}^{n} (W_j \times X_{ij}) \]  

(7)

According to the above steps, the weights of the tertiary indicators calculated by using Excel software are shown in Table 2. Then, the overall scores of science and technology achievement transformation capability and the scores of indicators at each level of universities in each province and city were calculated. The 31 provinces and cities were ranked according to the size of the scores. The results are displayed in Table 3.

5.2 Systematic Clustering Analysis

In order to better assess the capability of universities in each region to transform scientific and technological achievements, this study further analyzed 31 provinces and cities using cluster analysis. SPSS software is used for analysis. S1, S2, S3 and S were selected as clustering variables. The clustering analysis was performed using the systematic clustering method, and a clustering dendrogram was derived. As shown in Figure 1.

6 RESULTS AND DISCUSSION

As can be seen from Table 3, Jiangsu Province ranks first in terms of overall score. This reflects the leading position of universities in Jiangsu Province in the transformation of scientific and technological outcomes. Beijing, Shanghai, Guangdong and Zhejiang follow Jiangsu Province in the ranking. The universities in these regions have relatively strong ability to transform scientific and technological achievements. Universities in the cities of Chongqing, Liaoning, Hunan, Anhui and Tianjin have a medium level of achievement transformation capability. However, Ningxia, Hainan, Qinghai and Tibet, provinces with
weaker economic bases, ranked low in the comprehensive score. This indicates that universities are less capable of transforming their achievements. Therefore, it can be seen that there are large gaps in the capability to transform scientific and technological achievements among universities in different regions.

### Table 3: Indicator score and ranking.

<table>
<thead>
<tr>
<th>Number</th>
<th>Province</th>
<th>S1</th>
<th>Ranking</th>
<th>S2</th>
<th>Ranking</th>
<th>S3</th>
<th>Ranking</th>
<th>S</th>
<th>Ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Beijing</td>
<td>0.3248</td>
<td>2</td>
<td>0.1648</td>
<td>4</td>
<td>0.1400</td>
<td>1</td>
<td>0.6385</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>Tianjin</td>
<td>0.0804</td>
<td>13</td>
<td>0.0865</td>
<td>14</td>
<td>0.0347</td>
<td>12</td>
<td>0.2015</td>
<td>14</td>
</tr>
<tr>
<td>3</td>
<td>Hebei</td>
<td>0.0706</td>
<td>17</td>
<td>0.0456</td>
<td>19</td>
<td>0.0087</td>
<td>18</td>
<td>0.1250</td>
<td>20</td>
</tr>
<tr>
<td>4</td>
<td>Shanxi</td>
<td>0.0481</td>
<td>22</td>
<td>0.0321</td>
<td>21</td>
<td>0.0054</td>
<td>20</td>
<td>0.0856</td>
<td>22</td>
</tr>
<tr>
<td>5</td>
<td>Neimeng</td>
<td>0.0124</td>
<td>26</td>
<td>0.0143</td>
<td>26</td>
<td>0.0012</td>
<td>27</td>
<td>0.0279</td>
<td>26</td>
</tr>
<tr>
<td>6</td>
<td>Liaoning</td>
<td>0.1118</td>
<td>10</td>
<td>0.1346</td>
<td>7</td>
<td>0.0331</td>
<td>13</td>
<td>0.2795</td>
<td>11</td>
</tr>
<tr>
<td>7</td>
<td>Jilin</td>
<td>0.0539</td>
<td>20</td>
<td>0.0420</td>
<td>20</td>
<td>0.0291</td>
<td>14</td>
<td>0.1250</td>
<td>19</td>
</tr>
<tr>
<td>8</td>
<td>Heilongjiang</td>
<td>0.0801</td>
<td>14</td>
<td>0.0579</td>
<td>18</td>
<td>0.0051</td>
<td>21</td>
<td>0.2015</td>
<td>14</td>
</tr>
<tr>
<td>9</td>
<td>Shanghai</td>
<td>0.2527</td>
<td>3</td>
<td>0.1068</td>
<td>9</td>
<td>0.1295</td>
<td>2</td>
<td>0.4891</td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>Jiangsu</td>
<td>0.3774</td>
<td>1</td>
<td>0.2674</td>
<td>1</td>
<td>0.1060</td>
<td>4</td>
<td>0.7508</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>Zhejiang</td>
<td>0.1551</td>
<td>8</td>
<td>0.0896</td>
<td>13</td>
<td>0.0988</td>
<td>5</td>
<td>0.3435</td>
<td>8</td>
</tr>
<tr>
<td>12</td>
<td>Henan</td>
<td>0.0873</td>
<td>12</td>
<td>0.0741</td>
<td>16</td>
<td>0.0149</td>
<td>16</td>
<td>0.1763</td>
<td>15</td>
</tr>
<tr>
<td>13</td>
<td>Hubei</td>
<td>0.1929</td>
<td>5</td>
<td>0.1053</td>
<td>10</td>
<td>0.0506</td>
<td>10</td>
<td>0.3488</td>
<td>7</td>
</tr>
<tr>
<td>14</td>
<td>Hunan</td>
<td>0.1229</td>
<td>9</td>
<td>0.0729</td>
<td>17</td>
<td>0.0602</td>
<td>9</td>
<td>0.2559</td>
<td>12</td>
</tr>
<tr>
<td>15</td>
<td>Guangdong</td>
<td>0.2199</td>
<td>4</td>
<td>0.1015</td>
<td>11</td>
<td>0.1264</td>
<td>3</td>
<td>0.4478</td>
<td>4</td>
</tr>
<tr>
<td>16</td>
<td>Guangxi</td>
<td>0.0541</td>
<td>19</td>
<td>0.0831</td>
<td>15</td>
<td>0.0040</td>
<td>22</td>
<td>0.1413</td>
<td>18</td>
</tr>
<tr>
<td>17</td>
<td>Hainan</td>
<td>0.0083</td>
<td>28</td>
<td>0.0054</td>
<td>29</td>
<td>0.0009</td>
<td>29</td>
<td>0.0145</td>
<td>29</td>
</tr>
<tr>
<td>18</td>
<td>Chongqing</td>
<td>0.0791</td>
<td>15</td>
<td>0.1721</td>
<td>2</td>
<td>0.0383</td>
<td>11</td>
<td>0.2895</td>
<td>10</td>
</tr>
<tr>
<td>19</td>
<td>Sichuan</td>
<td>0.1051</td>
<td>11</td>
<td>0.1512</td>
<td>5</td>
<td>0.0724</td>
<td>6</td>
<td>0.3286</td>
<td>9</td>
</tr>
<tr>
<td>20</td>
<td>Guizhou</td>
<td>0.0204</td>
<td>24</td>
<td>0.0154</td>
<td>25</td>
<td>0.0027</td>
<td>23</td>
<td>0.0385</td>
<td>25</td>
</tr>
<tr>
<td>21</td>
<td>Yunnan</td>
<td>0.0280</td>
<td>23</td>
<td>0.0254</td>
<td>23</td>
<td>0.0009</td>
<td>28</td>
<td>0.0543</td>
<td>23</td>
</tr>
<tr>
<td>22</td>
<td>Xizang</td>
<td>0.0000</td>
<td>31</td>
<td>0.0002</td>
<td>31</td>
<td>0.0001</td>
<td>30</td>
<td>0.0003</td>
<td>31</td>
</tr>
<tr>
<td>23</td>
<td>Shanxi</td>
<td>0.1781</td>
<td>6</td>
<td>0.1397</td>
<td>6</td>
<td>0.0692</td>
<td>7</td>
<td>0.3869</td>
<td>6</td>
</tr>
<tr>
<td>24</td>
<td>Gansu</td>
<td>0.0187</td>
<td>25</td>
<td>0.0237</td>
<td>24</td>
<td>0.0015</td>
<td>26</td>
<td>0.0439</td>
<td>24</td>
</tr>
<tr>
<td>25</td>
<td>Qinghai</td>
<td>0.0048</td>
<td>30</td>
<td>0.0026</td>
<td>30</td>
<td>0.0000</td>
<td>31</td>
<td>0.0074</td>
<td>30</td>
</tr>
<tr>
<td>26</td>
<td>Ningxia</td>
<td>0.0066</td>
<td>29</td>
<td>0.0072</td>
<td>28</td>
<td>0.0019</td>
<td>25</td>
<td>0.0156</td>
<td>28</td>
</tr>
<tr>
<td>27</td>
<td>Xinjiang</td>
<td>0.0121</td>
<td>27</td>
<td>0.0095</td>
<td>27</td>
<td>0.0025</td>
<td>24</td>
<td>0.0241</td>
<td>27</td>
</tr>
</tbody>
</table>

Note: Score for the level of science and technology research and development, S1; Score for the level of scientific and technological output, S2; Score for the level of transformation of scientific and technological achievements, S3; Overall Score, S

### Table 4: Clustering results.

<table>
<thead>
<tr>
<th>Category</th>
<th>Province</th>
<th>S1 Average score</th>
<th>S2 Average score</th>
<th>S3 Average score</th>
<th>S Average score</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Beijing, Jiangsu</td>
<td>0.35</td>
<td>0.22</td>
<td>0.13</td>
<td>0.69</td>
</tr>
<tr>
<td>II</td>
<td>Shanghai, Guangdong, Liaoning, Chongqing, Sichuan, Zhejiang, Shanxi, Shandong, Hubei</td>
<td>0.16</td>
<td>0.13</td>
<td>0.08</td>
<td>0.37</td>
</tr>
<tr>
<td>III</td>
<td>Tianjin, Hebei, Shanxi, Neimenggu, Jilin, Heilongjiang, Anhui, Fujian, Jiangxi, Henan, Hunan, Guangxi, Hainan, Guizhou, Yunnan, Xizang, Gansu, Qinghai, Ningxia, Xinjiang</td>
<td>0.04</td>
<td>0.04</td>
<td>0.01</td>
<td>0.10</td>
</tr>
</tbody>
</table>
Further analysis illustrates that there are some provinces and cities with large differences in the ranking of each level of indicators. The universities in two provinces, Heilongjiang and Hubei, have higher levels of scientific and technological research and development process than their achievements output and transformation process. In Liaoning, Anhui, Fujian, Guangxi and Chongqing, the level of scientific and technological output of universities is stronger than their own level of research and development and transformation of results. The universities in Jilin, Jiangxi and Shandong are stronger in the third stage than in the first two stages. Therefore, universities in some regions have uneven development of their levels in each stage of the whole process of science and technology achievement transformation.

As can be seen from Figure 1, the 31 provinces and cities in China can be divided into 3 categories, defined as I, II and III. The specific classification is shown in Table 4. Table 4 shows that Beijing and Jiangsu belong to category I. The universities in this category have the highest average S scores, which indicates that the universities in Beijing and Jiangsu have the strongest capacity to transform scientific and technological achievements among the universities in China. The high level of economic development and the strong government investment in scientific research in Beijing and Jiangsu provide good environment and conditions for the transformation of scientific and technological achievements. In addition, these provinces have many research institutes, colleges and universities as well as scientific research talents, which strongly support the transformation of science and technology productions of universities.

Shanghai, Guangdong, Liaoning, Chongqing, Sichuan, Zhejiang, Shanxi, Shandong and Hubei belong to category II. Universities in category II have relatively strong ability to transform scientific and technological achievements. It can be seen that most regions in this category have average economic strength. Universities have uncoordinated development of strength in each stage of science and technology achievement transformation activities. Thus, universities should focus on strengthening the management of science and technology achievement transformation and planning the whole process of achievement transformation in an integrated manner.

Other 20 provinces and cities such as Tianjin, Hebei, Shanxi belong to category III. In these provinces and cities, universities have the capacity to transform scientific and technological achievements at the medium level and below. Compared with other categories, universities have more balanced development strength in all stages. Most of the provinces and cities in this category belong to the central and western regions of China with weak economic power. Coupled with the low government investment in scientific research, the number of scientific and technological achievements produced is relatively small. In addition, the market development in these provinces and cities is not perfect and the competition environment in the market is not mature enough. These lead to the scientific and technological achievements cannot be popularized and applied, which in turn restricts the transformation of scientific and technological achievements of universities. Therefore, higher education institutions in these regions should introduce social funds, technical support and scientific and technological talents from multiple channels. The government should increase support for scientific research, improve relevant policies and make efforts to develop the market.
7 CONCLUSIONS

Based on the whole process of transformation of scientific and technological achievements, this study established an assessment index system for the transfer capability of scientific and technological achievements in universities from three dimensions of research and development, output and transfer of scientific and technological achievements. A comprehensive evaluation of 31 provinces, cities and autonomous regions across China was conducted using entropy weighting and clustering methods. The research results indicate that the transformation capability of scientific and technological achievements of universities in different regions of China varies widely. In addition, universities in some regions have uneven development in the level of each stage of the conversion process. In this study, 31 provinces and cities were divided into three categories based on the scientific and technological achievement transformation capability of universities in each province and city. The provinces in category I have high economic level. Local governments provide sufficient financial support for the transformation of scientific and technological achievements. Talents gather here, and universities have a strong ability to transform their achievements. The provinces in category II have an average economic base. The universities in these regions have uneven development of strength in each stage of conversion of science and technology achievements. Category III has a weak economic foundation and little support from local government for scientific research. These regions are relatively short of scientific and technological talents. There is a lot of room to promote the capability of higher education institutions to transform their achievements.

7.1 Theoretical and Practical Implications

This research complemented and improved the evaluation research on the transformation of scientific and technological achievements and enriches the theoretical understanding of the science and technology achievement transformation. Thus, this research has certain theoretical value significance. This research carefully analyzed the empirical results, and provided an empirical basis for scientifically and accurately judging the scientific and technological achievement conversion ability of universities in each province and city. In addition, present study analyzed problems of universities in the process of transformation of scientific and technological achievements through the empirical results, and pointed out the path direction for universities to improve the transformation ability of scientific and technological achievements. Therefore, present study also has certain practical significance.

7.2 Limitations and Future Research

Firstly, this research evaluated the transformation capability of scientific and technological achievements of universities in each province and city in China from a horizontal perspective. Future researches can carry out relevant studies from the longitudinal timeline in order to understand the dynamic trend of the development of science and technology achievements conversion capacity of universities in various regions of China over the years. Secondly, considering the feasibility and data collectability, only some more representative and operational tertiary indicators were selected in this study. More indicators can be included in future studies for a more comprehensive and specific analysis.
Acknowledgements: This study was supported and funded by the Key R & D Plan of Shandong Province (Soft Science Project) (2021RZB01002).

REFERENCES

Big Data Analysis of the Business Model Innovation

Yanjing Gu
Corresponding author: Yanjing Gu 40520108@qq.com
Fuzhou Melbourne Polytechnic, China

Abstract: As the swift technological development, it seems that the speedy changes appears in the current society. Enterprises face severe competitions, not only in products and services innovation, but also in business model innovation. In particular, an increasing amount of attention had been paid on business model innovation over the last one decade. In this way, along with the progress of data economy, big data obviously become the core assets in the business competition, such as artificial intelligent, internet-of-things, etc. This paper aims to introduce the impact of data-driven business model innovation. At the same time, based on the combination of the big data and business model innovation, Freshippo will be chosen as an example to illustrate the huge impact on the application of big data in the business model innovation. Its new online shopping format affect the business model of supermarket in the retail industry.

Keywords: Business Model Innovation, Big Data, Shopping Format.

1 INTRODUCTION

Zott & Massa (2011) mentioned that business model can be regarded as a structural template, which has an influence on the macro-management research. Meanwhile, Teece (2010) revealed that business model can be understood as a structure of the value chain, which can be connected with competitive strategy. Along with the development of digital economy, e-commerce and technology management play an essential role in current society [11]. In detail, a growing number of data can be recorded by the application of technology instead of recording it by hand. Apparently, it shortens the time of collecting data and minimize the errors of recording, which promotes the efficiency and effectiveness. Also, as the related digital technologies are applied in commercial market, e-commerce and commercial intelligent data can provide necessary information for customers. That is to say, big data can create huge commercial opportunities. However, in practice, many enterprises still face much challenges as the business model innovation, especially to the small-size or medium-size companies.

In order to adapt the digital transformation, how should the firm to exploit the business model innovation is a competitive issue that should be taken into consideration. Evans et al (2017) mentioned that business model innovation aim to enhance the market competition to add value of the products or services. In the digital economy area, entrepreneurs can take advantage of the advanced technological methods to realign their business model. Innovative change in the business model may lead to a performance outcomes.
2 DISCUSSION THE DATA-DRIVEN BUSINESS MODEL INNOVATION

2.1 Conceptual Framework of Business Model Innovation

Business model innovation can be viewed as a process of collecting, organizing and summarizing internal and external data, which can acquire the market expectation of the products, product recommendation and the identification of customer need [1]. Moreover, Wirtz, Pistoia, Ullrich and Gottel (2016) represented that many previous scholars contributed on the research of business model innovation. In fact, Amazon, Alibaba, Facebook, etc, some of the famous companies already have financed huge amount of funds in capital market. Based on the successful business model innovation, a significant competitive advantages can benefit the development of the enterprises. Obviously, the technological progress is the main component that encourage the application of business innovation.

Concerning the mainly factors that affect the business model innovation, value creation, value delivery and value capture are the mainly essential aspects. Specially, value creation can be defined as the re-engineering of the core assets and skills to create an extra values [4]. According to the combination of the data, calculation and technology, the format of value creation can be reflected as the business model innovation. By the different categories of customers, big data analysis can provide more potential business opportunities. Regarding to the value delivery, the methods and approaches should embody a significant change during the process of the business model innovation. It associates with the internal and external ecosystems. In the new commercial model, promoted innovation business steps and operating activities should be applied. Companies have to link the business model with the digital platform to enhance the efficiency to delivery the business value. Lastly, value capture represents that how can the company acquire the capacity to create the value. The acquisition of the business value constitutes with the source of the revenue, the control of costs and the creation of business value. As the reset of the commercial steps in the business activities, it promotes the companies to adjust their business model to meet the market demand. However, Giesen et al. (2007) stated that business model innovation disrupted the supply chain, business operating model and business organizations. In this context, it fails to yield significant financial performance among the various business model.
2.2 Building Business Models by Big Data

Big data can be regarded as a competitive benefit and a forecasting for the successfully business model. In detail, the most essential characteristic of big data includes volume, velocity and variety. By utilizing the high-volume data, it is clear for decision-maker to summarize the trends of market and the customer base. Also, the main purpose of analyzing high-dimensional data is to make future prediction. Based on the different size and scope of the business, the usage of the big data analysis can justify the suitable way of business running. At the same time, with the digital pace of change, business activities move away from the conventional models to more technological models, which boosts the progress in e-commerce and the design of business strategies.

Notability, current market is filled with a large amount of data, which ranges from the individuals, companies and national government. In this way, companies should create appropriate business model that matches the customer needs by contributing on the relevant external data. As the data shown in Table 2, it forecasts that the China Fresh E-commerce Market Size can reach 419.83 billions in 2023. It represents a significant increase during the last 7 years by the rapid development of digital technology. Moreover, according to the assortment data from many sources, a complex analysis can be conducted to create revenues for the company. The effective methods to control the costs should also be considered. By extracting the big volume data, it enables the company to capture the business opportunities and build data-based software to deal with the enterprises problems.
From the Internet technology perspective, artificial intelligence, big data, internet-of-thing is widely applied in China fresh market. By online delivering system, the best delivery routes and districts can be matched by the customers’ orders. It can improve the delivery efficiency to large extent.

In order to illustrate the consumers’ willingness of online shopping, 1650 potential customers were selected to make this survey. According to the data from Table3, it indicates that 86.7% people are satisfied and very satisfied with deploying the e-commerce platform to purchase their groceries in the daily life. At the same time, as the pandemic outbreaks, fresh food e-commerce platform users increased dramatically. More fresh e-commerce platforms emerged in the market, like Daily youxian, Dingdong Maicai, Meituan, etc. In this way, the market size of China’s fresh food e-commerce industry maintained a strong market position, compared with the conventional fresh food industry.

Freshhippo is an example to illustrate the data-driven business model innovation in Chinese market. Freshhippo is established in Mar. 2015, Freshhippo is subordinated to Alibaba, which is a data-driven platform in the digital economy area. Meanwhile, Freshhippo is a completely a new retail format of offline supermarkets. Until to Mar. 2021, there are more than 220 stores all around China. According to the untouched delivery, the pandemic of covid-19 boost the accelerated development of this untouched delivery format. Based on the sales data of Freshhippo in 2021, it indicates that the sales revenue online is 30 percent more than the sale
revenues in 2020. That is to say, a growing number of consumers tend to choose online shopping. In this paper, a sophisticated analysis will be made to research the change of data-driven business model innovation.

![Willingness Chart](image)

**Table 3:** China Fresh E-commerce Consumer Introduction Willingness Survey in 2022

### 3.1 Technological Innovation Create Business Value

Related to the application of the online digital shopping pattern, customers can select, order the commodities and using mobile payment by Freshippo app. When the Freshippo store reaches the customer order, shopper will start to prepare the commodities and delivery to customers as soon as possible. Freshippo takes advantage of the variety of relevant sources, promotes the innovative business model, such as combing online format with offline format, positioning accurate customers, etc. In this way, the digital innovative business model can upgrade the company’s market share.

Concerning the application of mobile internet, e-cloud technology, big data analysis, more enterprises prefer to choose O2O business model. As the special format of the supermarkets, Freshippo reconstruct the format by connecting the online shopping format with the offline format to satisfy the customer need and enhance the satisfaction of various customers. This reconstruction of the shopping format transfers the conventional business patterns. In this way, this advanced business pattern create entirely curious shopping experiences for customers by the new digital technology. Apart from this, this updated format break the limitation of time and space. Customers can shop around their workplace or in their apartment. It directly broadens the customers set and facilitate the customers. Hence, using the newer business pattern can promote the creation of economic value and increase the value of the enterprise.
3.2 Data-Driven Analysis in the Business Model Dimension

Due to the shopping recording online by using Freshippo App, the internet technology can be utilized to analyze the customers behaviour, like big data analysis, cloud calculation, etc. By assorting the customers information and cloud storage, Massa and Tucci (2013) noted that it gradually transfer the business creation model and create the value added way to sell existing product. At the same time, software platform can make the best of big data and classify the target customers based on the current market situation. It can deliver the precise information to the potential customers to large extent, which differs a lot from the offline supermarkets. In this way, more efficiency information that customers tend to interest can be received.

Interesting, Freshippo has their own delivery services. Thus, it is unnecessary for customers to worry about the high volume online orders, even though the orders outbreak. In this way, Freshippo created the Self-Pick’s Go service. According to the online orders, shoppers can prepare and distribute the food into the self-pick box. This service facilitates the customers around the stores to enjoy their fresh food without any contact or risks. From the Freshippo’s perspective, it directly cut the labor costs. They can collect their food in the self-pick box by scanning the codes, it improves the services efficiency and brings convenient for the commuters during the peak hours. Moreover, Freshippo also provides Freshippo F2 and Freshippo mini formats. The new concept format of supermarket incurs the interest of different aged person, which addresses the customers’ needs.

4 CONCLUSION

Digital innovation already become an inevitable part of the business model innovation, it can promote the value creation, value delivery and value capture of enterprises. Based on the application of internet, cloud technology and big data, the digital software platform can be deployed for various customers. As the support of the advanced digital technology, the combination of online and offline shopping patterns can promote the customer satisfaction and create higher brand value. Meanwhile, in the rapid digital area, business model can be vital factor for the sustainable development. That means the successful business model can enhance the sale revenues and increase the competition in the fierce business environment. Also, compared with the conventional business model, digital innovative business model can be more flexible and variety.

REFERENCES

How Shareholders' Right of Speech Affects the Performance of Green Innovation--Big Data Practice Based on R Language Brucer Package

Yulu Chen¹,a, Junzheng Wu¹,b, ChuHan Lin¹,c

1206480228@qq.coma, 1259934507@qq.comb, 2191573215@qq.comc

Strait College of Minjiang University, Minhou County, Fuzhou City, Fujian Province, China¹

Abstract: Objective: In the internal governance of enterprises, the impact of shareholders' voice on the performance of green innovation is worth empirical research. Process: This paper takes a large number of data of listed companies from 2010 to 2020 as the research object, uses 28940 data, uses R studio environment to build a time fixed effect model, and uses OLS and other methods to the relationship between shareholder voice and corporate green innovation performance. Conclusion: The improvement of shareholders' discourse power is not conducive to improving the green innovation output performance of enterprises; Shareholders' discourse power affects the green innovation performance of enterprises through its role in enterprise growth. The growth of enterprises is conducive to improving the green innovation performance of enterprises. After a series of robustness tests, the conclusion is still valid.

Keywords: Shareholders' Right Of Speech, Green Innovation Performance, Enterprise Growth, Equity Concentration.

1 INTRODUCTION

In the R language environment, the traditional analysis of economic big data is more the OLS linear regression in the application base package. However, when exporting the results, it is often unable to obtain good visualization results. In actual operation, you will use different software to analyze mediation/regulation effects - Mplus, LISREL, the PROCESS plug-in of SPSS, etc. Among them, the most popular is probably the PROCESS macro for SPSS and SAS, because it is simple to operate, quick to use, and has a wide variety of models. Of course, if you want to conduct more complex path analysis and latent variable modeling, you still need to use Mplus. The bruceR:: PROCESS() function uses the mediation package for mediation effect analysis, the interactions package for simple slope analysis, and the lavaan package for chain multiple mediation analysis. The mediation with adjustment calls both mediation and interactions packages.

BruceR:: PROCESS() function does not refer to the official script provided by Hayes (after all, there are 5300 lines of code officially), but strict comparison is made during development to ensure the accuracy of the results. This paper will use bruceR package to analyze the real big data in China.
This requirement highlights the country's emphasis on innovation and R&D from the macro level. As one of the key subjects to implement the government's policy objective of "green innovation performance", listed companies take business innovation and technological change as their core competitiveness. Green innovation performance is an important part of their capital destination and plays a decisive role in improving the economic efficiency of enterprises. In recent years, as "innovation is the first productive force" and "the concept of sustainable development" are increasingly valued by enterprises, the research on green innovation performance and enterprise development has become the research direction of many scholars. However, previous studies have paid more attention to the relationship between green innovation performance and operating profit, profitability and innovation performance, while the research on internal factors of green innovation performance of enterprises is less. At the same time, due to the concentration of equity in some listed companies, shareholders with large equity shares often make biased decisions in order to safeguard their own interests, which affect the level of green innovation performance of enterprises, or affect market performance by changing green innovation performance to achieve profits. Based on the above analysis, this paper selects relevant data from 2010 to 2020 to explore the relationship between equity concentration, enterprise growth and green innovation performance of listed companies through theoretical analysis and empirical testing, focusing on the intermediary role of enterprise growth in the process of equity concentration affecting market performance, with a view to providing strong empirical evidence for the research on influencing factors of market performance, And put forward targeted policy recommendations. At the same time, there are many researches on the external factors that affect the performance of green innovation in the current academic circle, but there are few researches on the internal factors of green innovation in enterprises. This paper will call bruceR package to study the internal factors that affect the green innovation of enterprises.

2 LITERATURE REVIEW AND RESEARCH HYPOTHESIS

2.1 Equity Concentration and Green Innovation Performance

Equity concentration is often used as a key indicator to reflect the equity distribution of a company. Some scholars believe that excessive equity concentration will inhibit the efficiency of green innovation performance transformation. Some scholars found that ownership concentration has a significant negative impact on green innovation performance. Many scholars have studied the regulatory effect of equity concentration, and some scholars have found that equity concentration negatively regulates the impact of equity incentive intensity on R&D expenditure. Some scholars also found that ownership concentration has no significant direct effect on R&D investment, but it plays a significant positive regulatory effect in the process of manager autonomy positively affecting R&D investment intensity. Innovative R&D activities are inherently risky, and the benefit of enterprise input and output is faced with greater uncertainty, which will affect the business performance and total profit of the enterprise, and ultimately determine the amount of dividends for equity investors. With the improvement of equity concentration, the higher the risk of enterprise technological innovation undertaken by major shareholders, the more they tend to operate conservatively, maintain the current business situation of the enterprise, and are unwilling to carry out innovation and change and invest in...
research and development activities. Some scholars pointed out that the investor protection environment in China is not perfect. Under the imperfect weak investor protection environment, large shareholders are likely to increase the value of their own corporate resources through other more convenient ways, such as maximizing private benefits through a large number of related party transactions, which is "more efficient" than the benefits obtained by investing in research and development activities. That is to say, the higher equity concentration restrains the R&D expenditure of enterprises. Some scholars believe that equity concentration will weaken the green innovation performance of enterprises [5]. Some people also get the conclusion that the high concentration of equity will inhibit the performance of green innovation through analysis. R&D activities are characterized by high investment and high risk, and the time span is relatively long [6]. Driven by the pursuit of short-term income goals and risk aversion, major shareholders are reluctant to carry out R&D activities. Based on the above analysis, this paper proposes hypothesis 1:

H1: Equity concentration has a negative effect on corporate green innovation performance.

2.2 Equity Concentration and Enterprise Growth

The smaller the equity concentration of the company, the lower the noise of the relevant variables used in the internal contract, and the more accurate and reliable they are, the more they will provide good incentive and supervision signals to the board of directors and senior management, so as to improve the growth of the enterprise [7]. Effective internal control can improve the accuracy of internal management reports or internal data information by reducing unintentional accounting errors or intentional accounting manipulation, inhibit the earnings management problems of the company, and enhance the competitive advantage of the enterprise [8]. One of the functions of low equity concentration is to supervise and feedback organizational activities. With the daily activities of the enterprise running together, equity concentration can guide various departments, provide products that satisfy customers, and enable the enterprise to maintain a better competitive advantage. Optimizing the internal environment of the enterprise and strengthening the construction and implementation of risk assessment can improve the internal communication status of the organization and improve the growth of the enterprise. Therefore, this paper proposes the following assumptions:

H2: Equity concentration is not conducive to the growth of enterprises.

2.3 Enterprise Growth and Enterprise Green Innovation Performance

Generally speaking, Chinese enterprises have such defects as high survival pressure, poor resource base and strong financing constraints, which often lead to insufficient motivation for enterprises to carry out green innovation. However, for high growth enterprises, green innovation is the main way to obtain competitive advantages and maintain their own long-term development. Especially in the case of increasingly fierce external competition, enterprises also have enough motivation to carry out green innovation to ensure their growth ability [9]. To be specific, first of all, green innovation means that the enterprise has obtained new innovation impetus. Even if the enterprise decides to invest in innovative projects with high uncertainty, it shows that they have the potential of marketization and can bring potential competitive advantages to the enterprise; Secondly, with the increasingly frequent international and domestic exchanges of knowledge and technology, the innovation support that enterprises can
obtain has been greatly improved compared with the past. Many enterprises can grow rapidly and become listed companies in a short time, which is based on the rapid flow of innovation resources; Thirdly, the improvement of education level has made the quality of technical innovation personnel generally improve. The promotion of innovation and entrepreneurship at the national policy level has made many graduates who have received better education in colleges and universities and researchers in major scientific research institutes have a strong combination of industry, education and research, which has made the innovation level of all kinds of enterprises can be improved as a whole. Based on the above theoretical analysis, this paper proposes the following research assumptions:

H3: Enterprise growth has a positive effect on green innovation performance, and it plays an intermediary role in the relationship between ownership concentration and green innovation performance.

3 RESEARCH DESIGN

3.1 Sample Selection and Data Source

This paper selects the basic data and green innovation data of listed companies in the manufacturing industry of Shanghai and Shenzhen Stock Exchanges from 2010 to 2020 as samples. CSMAR database is the data source of this study, and data processing is conducted through R studio software.

At the same time, this paper has carried out the following processing: 1% and 99% percentile winsorize for all continuous variables to eliminate the influence of abnormal values; Some missing data shall be supplemented appropriately. These data are missing because the missing value itself represents that the item is 0, so the missing value is added as 0; Some industries, such as financial industries, which are not suitable for measuring the innovation ability of enterprises with the number of patents, have been eliminated. According to the above standards, 28940 observations were finally obtained in this paper.

3.2 Model Building

In this paper, the fixed effect model is used to empirically demonstrate the fixed effect model, namely the fixed effect regression model (FEM for short), which is a panel data analysis method. It refers to the experimental design that the experimental results only want to compare the differences between specific categories or categories of each autovariable and their interaction effects with specific categories or categories of other autovariables, and do not want to infer from this to other categories or categories not included in the same autovariable. Fixed effect regression is a variable method that varies with individuals but not with time in spatial panel data.

In order to test the negative effect of H1 ownership concentration on enterprise green innovation performance, this paper constructs the following model 1:

$$GI = \alpha + \beta_1 FSR + \gamma Controls + \epsilon.$$ (1)
In order to test the mediating role of H2 and H3 enterprise growth in equity concentration on enterprise green innovation performance, this paper constructs the following models 2 and 3:

\[ GI = \alpha + \beta_1 GROW + \gamma Controls + \varepsilon \] (2)

\[ GI = \alpha + \beta_1 GROW + \beta_2 FSR + \gamma Controls + \varepsilon \] (3)

3.3 Variable Definition

3.3.1 Interpreted variable

In order to measure the green innovation performance of enterprises, this paper selects the number of green patents measuring the green innovation achievements of enterprises in the CSMAR database as the green innovation performance of enterprises.

3.3.2 Explanatory variable

This paper chooses the largest ownership concentration as a substitute variable for the concentration of corporate shareholders’ discourse power.

3.3.3 Intermediary variable

This paper uses the asset to debt ratio in the enterprise financial data to quantify the alternative variables of enterprise growth.

3.3.4 Control variable

According to the experience of previous research, in order to control the impact of different enterprise characteristics on innovation output, this paper uses enterprise scale; Book to market ratio; Asset liability ratio and total asset turnover ratio are the control variables of this model.

3.4 Code

In the R language environment, the bruceR code is called and the adjustment effect is checked as follows:

```r
# Intermediary Analysis
PROCESS(total, y="grow", x="RD", meds = "fsr",
       covs = c("ROA","dar","mcr","size"),
       ci="boot",nsim = 1000,seed = 1)

# Robustness Check
PROCESS(total, y="RD", x="tsr", meds = "grow",
       covs = c("ROA","dar","mcr","size"),
       ci="boot",nsim = 1000,seed = 1)
```
4 EMPIRICAL RESULTS

4.1 Regression Results

First of all, this paper uses the fixed effect model for empirical estimation. Columns (1) and (2) of Table 1 respectively examine the impact of the implementation of shareholders’ discourse and enterprise growth on the innovation output and innovation efficiency of enterprises. It can be seen from Table 3 (1) that the coefficient of fsr is -0.086, which is significantly positive at the level of 5%, indicating that the implementation of the concentration of shareholders' right of speech significantly inhibits the total number of green patents of enterprises. In (3), the coefficient of growth is -0.393, which is not significant, indicating that the growth of enterprises presents a negative full intermediary role. Therefore, the assumption is supported by data.

<table>
<thead>
<tr>
<th></th>
<th>(1) RD</th>
<th>(2) grow</th>
<th>(3) RD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>10.348 (12.590)</td>
<td>33.467 *** (0.082)</td>
<td>10.348 (12.591)</td>
</tr>
<tr>
<td>ROA</td>
<td>-62.207 (41.050)</td>
<td>6.573 *** (0.268)</td>
<td>-59.627 (41.476)</td>
</tr>
<tr>
<td>mcr</td>
<td>-15.638 (60.581)</td>
<td>5.340 *** (0.395)</td>
<td>-13.542 (60.773)</td>
</tr>
<tr>
<td>Size</td>
<td>-8.744 (12.006)</td>
<td>1.585 *** (0.078)</td>
<td>-8.122 (12.091)</td>
</tr>
<tr>
<td>fsr</td>
<td>-0.086*** (0.008)</td>
<td>-0.000 *** (0.000)</td>
<td>-0.000 (0.008)</td>
</tr>
<tr>
<td>Grow</td>
<td></td>
<td></td>
<td>-0.393 (0.902)</td>
</tr>
<tr>
<td>R^2</td>
<td>0.000</td>
<td>0.046</td>
<td>0.000</td>
</tr>
<tr>
<td>Adj. R^2</td>
<td>0.000</td>
<td>0.046</td>
<td>0.000</td>
</tr>
<tr>
<td>Num. obs.</td>
<td>28940</td>
<td>28940</td>
<td>28940</td>
</tr>
</tbody>
</table>

Note. * p < .05, ** p < .01, *** p < .001.

4.2 Robustness Check

In this paper, the method of replacing explanatory variables is used to test the robustness. This paper changes the quantitative method of shareholders’ voice from the shareholding ratio of the largest shareholder to the shareholding ratio of the top ten shareholders. The results are as follows, and the assumption is valid, indicating that the conclusion is robust.

<table>
<thead>
<tr>
<th></th>
<th>(1) RD</th>
<th>(2) grow</th>
<th>(3) RD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>250.784 *** (8.802)</td>
<td>10.348 (12.590)</td>
<td>250.784*** (8.802)</td>
</tr>
<tr>
<td>ROA</td>
<td>86.948 ** (28.945)</td>
<td>-54.590 (41.400)</td>
<td>86.944 ** (28.946)</td>
</tr>
<tr>
<td>mcr</td>
<td>-68.038 (43.195)</td>
<td>158.405 * (61.782)</td>
<td>-68.027 (43.201)</td>
</tr>
<tr>
<td>Size</td>
<td>-369.271 ***</td>
<td>-6.233 (60.865)</td>
<td>369.271***</td>
</tr>
</tbody>
</table>
5 CONCLUSIONS

The report of the 20th National Congress of the Communist Party of China has repeatedly emphasized that we should further promote the energy transformation and adhere to the core position of innovation in China's overall modernization drive. In this macro context, there are obvious differences in existing research on the internal relationship between shareholders' voice and green innovation of enterprises. The academic and practical circles have not reached a consensus on whether the shareholders' right of discourse has aggravated the exclusion of internal resources from green innovation or promoted the integration of internal resources and green innovation. Therefore, this paper conducts research on the impact of shareholders' discourse power on green innovation of enterprises, and empirically tests the impact and mechanism of shareholders' discourse power on green innovation of enterprises using fixed effect model and intermediary effect model. The research conclusions are as follows: First, shareholders' discourse power inhibits green innovation of enterprises. After excluding the interference of endogenous problems and a series of robustness tests, the conclusion is still robust. Secondly, further research found that shareholders' right of discourse inhibits the performance of green innovation by reducing the growth of enterprises.

6 POLICY SUGGESTIONS

Based on the above research, this paper puts forward the following suggestions:

1. Controlling shareholders' equity concentration is an important part of deepening the reform of listed enterprises, which should give full play to the characteristics and advantages of the integration of multiple ownership capital, form collaborative innovation forces with small and medium-sized shareholders as the platform, improve the level of green innovation of enterprises, and achieve the organic unity of economic and social benefits;

2. Heterogeneity characteristics such as subordination level and functional positioning have an important impact on environmental technology R&D of listed enterprises. The introduction of small and medium-sized equity should fully consider the characteristics of enterprises themselves, and promote enterprise reform in a hierarchical and classified manner;

3. Environmental institutional pressure is an important driving force for green innovation of enterprises. External governance measures such as increasing environmental subsidies and increasing analysts' attention should be taken to foster a good external environment for green
innovation incentives and transparent information disclosure, and guide and encourage enterprises to conduct green innovation research and development.

REFERENCES

Dynamic Evaluation of Eco-Economic Development in Prefecture-Level Cities in Yunnan Province

Zhanni Huang1,a, Yunhe Zhao1,b*, Hansheng Hu1,c, Haoyang Shen1,d, Sirong Shang1,e

137570387@qq.coma, zyh961019@163.comb, 2371948464@qq.comc, shenhaoyang1019@icloud.comd, 3380511836@qq.come

School of Statistics and Mathematics, Yunnan University of Finance and Economics, No. 295 Longquan Road, Kunming, China1

Abstract: Based on the theory of ecological economics, ecological economy studies the sustainable development of natural socio-economic system. This paper constructs a dynamic evaluation system of ecological economy development, sets 18 secondary indicators from three perspectives of economic subsystem, social subsystem and ecological subsystem, and improves the entropy weight method by assigning weights to the indicators to calculate the evaluation scores of ecological economy development of eight prefecture-level cities in Yunnan Province. Through the analysis of panel evaluation data from 2010 to 2020, the following conclusions are drawn: (1) The overall development of the ecological economy of prefecture-level cities in Yunnan Province is good, with a large increase; (2) The prefecture-level cities in Yunnan Province as a whole have unbalanced development in the three subsystems of economy, society and ecology; (3) The development of all localities and cities in Yunnan Province is uneven, Kunming has a large development advantage, and Lijiang's ecological economic level is weak as a whole. In view of the above conclusions, this paper gives corresponding suggestions to help Yunnan Province's ecological economy to achieve further sustainable development.

Keywords: Ecological Economy, Development Evaluation, Entropy Weight Method, Panel Data, Sustainable Development.

1 INTRODUCTION

Under the background of China's socio-economic development increasingly pursuing quality and coordination, the sustainable development of ecological economy has attracted widespread attention [12]. In October 2022, General Secretary Xi Jinping pointed out in the report of the 20th National Congress of the Communist Party of China that the overall goal of China's development by 2035 includes: Build a modern economic system, form a new development pattern, basically realize the modernization of new agriculture, and widely form a green production and lifestyle, carbon emissions peak and decline steadily, and the ecological environment fundamentally improves.

Eco-economic evaluation analyzes the degree of coordination between the socio-economic and ecological ecosystems of different regions by incorporating the research elements of ecological economics [6]. The American economist Boulding (1966) [11] first proposed the concept of “ecological economics”, which is different from the previous development method of blindly...
developing the economy and ignoring the environment, he linked ecological factors with economic development. Subsequent researchers pointed out that research based on resource distribution is carried out from different perspectives, involving the fields of ecology, geography, management and economics. China's research on ecological economy is late, and Zhu Dajian (2009) proposed the difference between ecological economics and mainstream economics, including ecological natural resources into the utility function formula, emphasizing its significance to China's scientific development. Diao Shangdong et al. (2013) integrated the government's institutional construction into the evaluation index system of ecological civilization construction, reflecting the correlation between socio-economic development and government policies. Sun Buzhong et al. (2017) introduced time factors for dynamic evaluation and analysis, and established a composite system related to the length of observation. Ecological economy research mainly deals with the sustainable development of natural socio-economic systems, which can be summarized into two categories: (1) Quantitative measurement of sustainable ecological economic development [2, 6-7]; (2) Research on the integration model of ecological economic system [10, 12]. However, most of the domestic research on ecological economy focuses on the Yangtze River Economic Belt, and relatively few research on the ecological economy in southwest China. Yunnan Province has the positioning of “the vanguard of ecological civilization construction” and the needs of economic transformation [11], so studying the ecological economy of Yunnan Province is of great significance to the development of sustainable ecological economy in China.

Based on the existing research, this paper constructs an ecological economic development evaluation index system with economic subsystem, social subsystem and ecological subsystem as the primary indicators, and 18 indicators such as local general public budget revenue and expenditure, natural growth rate, and total carbon emissions as secondary indicators. At the same time, eight prefecture-level cities in Yunnan Province were selected as the research objects, and the sample data from 2010 to 2020 were selected to evaluate the dynamic development of ecological economy through the panel entropy weight method, and the data analysis results were analyzed and suggestions were made.

2 RESEARCH METHODS

2.1 Entropy weight method based on panel data improvement

The entropy weight method is a measurement method for assigning index values according to the amount of information provided by the observation values of each selected index, and as an objective weighting method, it has been applied to the study of ecological economy many times, and it has strong adaptability and can objectively evaluate the system.

This paper improves the panel data of three dimensions of city, time and indicators, and improves on the basis of the core idea of entropy weight method, and sets C city objects, i indicators and t time dimensions (time), through the standardization of data preprocessing, from the original two-dimensional data to form panel data indicators, and then carry out subsequent feature weight, entropy, weight, score calculation and analysis. The specific process of model building is as follows:
2.1.1 Standardization of panel indicators

For the standardization of indicators, it is necessary to achieve the three goals of positive indicators, dimensionless and normalization, so there are different operations for the processing of positive indicators and negative indicators.

Positive indicator standardization:

\[
y_{ij} = \frac{x_{ij} - x_{i \text{min}}}{x_{i \text{max}} - x_{i \text{min}}} \quad (1)
\]

Normalization of negative indicators:

\[
y_{ij} = \frac{x_{i \text{max}} - x_{ij}}{x_{i \text{max}} - x_{i \text{min}}} \quad (2)
\]

where represents the value of column \( y_{ij} \)th of row \( i \) in the normalized matrix. After standardizing the positive and negative indicators, the index data are displayed in the range of \((0,1)\).

Since the panel data involves three perspectives: city objects, indicators and time dimensions, it can be divided into double and two-dimensional data, and its indicators are standardized first, and the standardized matrix obtained, namely city-indicator matrix and \( Z_{ci} \) time-indicator matrix \( Z_{ti} \). At the same time, in the construction of panel indicators, the formula of formula (3) is improved, and the improved panel indicator standardization matrix is obtained \( Z_{cti} \).

\[
Z_{cti} = \sqrt{Z_{ci} \cdot Z_{ti}} \quad (3)
\]

2.1.2 Feature weight

Calculate the proportion of each indicator in the panel indicator to all indicators, that is, the feature weight expressed as the indicator \( Y_{cti} \).

\[
Y_{cti} = \frac{Z_{cti}}{\sum_{c=1}^{C} \sum_{t=1}^{T} Z_{cti}} \quad (4)
\]

2.1.3 Calculate the entropy value

Information entropy is a measure of the degree of information order, and the entropy value is calculated based on the above step feature weight \( e_t \).

\[
e_t = -\frac{1}{\ln(CT)} \sum_{c=1}^{C} \sum_{t=1}^{T} Y_{cti} \cdot \ln(Y_{cti}) \quad (5)
\]
2.1.4 Calculate the weight

Each indicator is objectively weighted after measuring the orderliness of data information.

\[
d_i = 1 - e_i 
\]

\[
w_i = \frac{d_i}{\sum d_i} 
\]

where \(d_i\) represents the information utility value and \(w_i\) represents the weight. The higher the degree of information order, the greater the information entropy, the less useful information provided, the smaller the information utility value, the smaller its weight, and vice versa.

2.1.5 Comprehensive scoring

According to the corresponding weight, the sample data of each indicator is weighted and summed, and the score and comprehensive score of each subsystem are calculated.

\[
S = \sum_{i=1}^{C} \sum_{j=1}^{I} w_{ij} \cdot Z_{ij} 
\]

2.2 Data Sources

The data used in this paper are mainly from China Urban Statistical Yearbook, Yunnan Statistical Yearbook, Scientific Data, and municipal statistical bureaus, and some missing data are supplemented by interpolation. Taking the prefecture-level cities in Yunnan Province as the research object, the statistical data of the eight prefecture-level cities involved were selected as data samples to ensure the validity of the data.

3 EMPIRICAL ANALYSIS

3.1 Overview of the study area

Yunnan Province is located in the Yunnan-Guizhou Plateau region of southwest China, shouldering the important responsibility of maintaining the ecology of southwest China, which is located in the transition zone of the ecosystem, with strong ecological environment heterogeneity and relatively fragile ecological system. Due to its geographical location, administrative background, resource conditions, and historical technology, the regional economic structure of Yunnan Province varies widely\(^{[11]}\).

Considering the feasibility of data collection, this paper selects eight representative prefecture-level cities (Kunming, Qujing, Yuxi, Baoshan, Zhaotong, Lijiang City, Pu'er City, Lincang City) for discussion.
Located in the north-central part of Yunnan Province, Kunming's economic development is relatively leading in the whole of Yunnan Province, providing perfect infrastructure and jobs, while continuously strengthening ecological governance.

Qujing City is located in the eastern part of Yunnan Province. It actively develops the economy, introduces leading enterprises in the industry, and coordinates the development of high-level ecological environmental protection and high-quality economy.

Yuxi City is located in the central part of Yunnan Province, with rich terrain and three-dimensional climate. The total economic volume has been continuously improved, the comprehensive strength has been steadily enhanced, and the implementation of ecological and environmental protection rectification and reform has been promoted.

Baoshan City is located in the southwest of Yunnan Province, rich in lignite reserves and geothermal resources. It has taken multiple measures at the same time, increased ecological protection, and built a solid ecological security barrier in the southwest.

Zhaotong City is located in the northeast of Yunnan Province, located in the transition zone from the Sichuan Basin to the Yunnan-Guizhou Plateau. Its gross economic output value has increased significantly, the comprehensive agricultural output value has developed rapidly, the ecological environment has achieved a good start, and the battle against pollution has achieved remarkable results.

Lijiang City is located in the northwest of Yunnan Province, at the intersection of Yunnan, Sichuan and Tibet provinces, with outstanding advantages in biological resources, water energy resources and tourism resources, and a wide variety of animals and plants, which is a key forest area in Yunnan Province Biodiversity areas and treasure trove of biological resources.

Pu'er City is located in the southwest of Yunnan Province, and its construction of the National Green Economy Experimental Demonstration Zone has a forest coverage rate of 74.59%.

Lincang City is located in the southwest of Yunnan, rich in water resources, and is an important national hydropower energy base.

### 3.2 Construction of indicator system and data sources

#### 3.2.1 Framework of evaluation index system

In order to accurately evaluate the regional ecological and economic level of prefecture-level cities in Yunnan Province, this paper mainly selects indicators based on the principle of systematicism, the principle of combining theory and practice, the principle of operability of indicators, and the principle of data availability. At the same time, combined with the existing research on ecological economy and the specific situation of various cities in Yunnan Province, this paper draws on He Weijun et al. (2016), Jia Haifa et al. (2020), Sun Buzhong et al. (2017), etc. According to the structural hierarchy analysis theory, the ecological economy evaluation index system of prefecture-level cities in Yunnan Province was constructed (Table 1), which took the economic subsystem, social subsystem and ecological subsystem as the first-level indicators. The Eighteen secondary indicators are selected based on the following:
(1) In the economic subsystem, the local general public budget revenue and expenditure, per capita regional GDP, and average wages of on-the-job employees reflect the degree of economic development, and the development of ecological economy is analyzed from the perspective of industrial structure of several industrial enterprises above designated size, the proportion of tertiary industry in GDP, and the total retail sales of social consumer goods.

(2) In the social subsystem, the total population and natural growth rate represent the number of people and the level of human capital, the number of registered unemployed in cities and towns, the number of employees in health and social security and social welfare industries reflect the level of social security, and the expenditure on education and science and technology represents the level of social education and quality.

(3) In the ecological subsystem, the centralized treatment rate of sewage treatment plants, industrial nitrogen oxide emissions, and total carbon emissions are used as indicators of waste emissions, while the proportion of urban construction land area, built-up area green coverage rate, and park green area are environmental protection construction indicators.

The above indicators can better reflect the actual situation of the ecological economy of prefecture-level cities in Yunnan Province, and provide real and reliable data for ecological economic evaluation and analysis, and the specific ecological economic evaluation index system is shown in Table 1 below. The weights obtained by entropy weighting method as shown in the table show that the economic subsystem (0.3355), social subsystem (0.3303), and ecological subsystem (0.3342) have equal status in ecological economic. The economic subsystem has a slightly higher weight, followed by the ecological subsystem, followed by the social subsystem.

3.3 Evaluation and analysis

According to the established dynamic evaluation index system of ecological economic development, the data information was analyzed by programming using the panel entropy weighting method with STATA software to obtain the individual secondary index scores, primary index (subsystem) scores and comprehensive scores of each region for each year. At the same time, the overall eco-economic development of Yunnan Province, the comprehensive score ranking of 8 prefecture-level cities and the dynamic situation of each region from 2010 to 2020 are discussed and analyzed.

<table>
<thead>
<tr>
<th>Evaluation indicators of eco-economic development</th>
<th>Level 1 indicators</th>
<th>Weight</th>
<th>Secondary indicators</th>
<th>Unit</th>
<th>Entropy</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Economic subsystem</td>
<td>0.3355</td>
<td></td>
<td>Local general public budget revenue and expenditure status of the whole city</td>
<td>million</td>
<td>0.0625</td>
<td>0.0570</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Average wages of on-the-job employees</td>
<td>Yuan</td>
<td>0.0931</td>
<td>0.0551</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Number of industrial enterprises above designated size</td>
<td>piece</td>
<td>0.0895</td>
<td>0.0553</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Total retail sales of consumer goods</td>
<td>million</td>
<td>0.0559</td>
<td>0.0574</td>
</tr>
</tbody>
</table>
### Table 2: Dynamic evaluation of ecological economic development of prefecture-level cities in Yunnan Province

<table>
<thead>
<tr>
<th>Year</th>
<th>Economic subsystem</th>
<th>Social subsystems</th>
<th>Ecological subsystems</th>
<th>Overall rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td>0.000606</td>
<td>0.001640</td>
<td>0.001445</td>
<td>0.003692</td>
</tr>
<tr>
<td>2011</td>
<td>0.000707</td>
<td>0.001708</td>
<td>0.001288</td>
<td>0.003704</td>
</tr>
<tr>
<td>2012</td>
<td>0.000941</td>
<td>0.001579</td>
<td>0.001407</td>
<td>0.003927</td>
</tr>
<tr>
<td>2013</td>
<td>0.001160</td>
<td>0.001962</td>
<td>0.001556</td>
<td>0.004678</td>
</tr>
<tr>
<td>2014</td>
<td>0.001217</td>
<td>0.001888</td>
<td>0.001663</td>
<td>0.004769</td>
</tr>
<tr>
<td>2015</td>
<td>0.001589</td>
<td>0.002098</td>
<td>0.001614</td>
<td>0.005301</td>
</tr>
<tr>
<td>2016</td>
<td>0.001890</td>
<td>0.002059</td>
<td>0.001901</td>
<td>0.005850</td>
</tr>
<tr>
<td>2017</td>
<td>0.001952</td>
<td>0.001984</td>
<td>0.001879</td>
<td>0.005815</td>
</tr>
</tbody>
</table>

**3.3.1 Dynamic evaluation and analysis of the overall ecological economic development of prefecture-level cities in Yunnan Province**

From the overall analysis, the comprehensive score of ecological economic development of prefecture-level cities in Yunnan Province showed an upward trend from 2010 to 2020, from 0.0037 to 0.0058. The increase was close to 57 percent, but in recent years the rate of growth has slowed and has shown a downward trend.
From the analysis of various subsystems, there is a significant increase in economy, society and ecology, of which the economic subsystem has the most obvious increase, close to 76%, the above data indicate in recent years, Yunnan Province has developed rapidly and played an important role in the development of ecological economy. The score of social subsystem has increased in previous years, and the score has decreased in recent years. The ecological subsystem as a whole is on an upward trend, but the increase rate is not very obvious, and there has been a downward trend in recent years. The index weights of the three subsystems of economy, society and ecology are almost the same, but there are obvious differences in the scores of the three subsystems in Yunnan Province, indicating that its ecological economic development has not maintained a balance in the three dimensions of economy, society and ecology, and has not achieved sustainable development in the field of ecological economy.

### 3.3.2 Dynamic Scoring Analysis of Ecological Economy in Cities at All Levels in Yunnan Province

From the perspective of prefecture-level cities, the comprehensive score ranking of 8 prefecture-level cities in Yunnan Province changes slightly every year, and there is an uneven development phenomenon in prefecture-level cities. The results of the dynamic eco-economic scoring of cities at all levels in Yunnan Province are detailed in the Table 3.

**Table 3: Ecological economy score and ranking of cities at the local level**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Kunming City</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Qujing City</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Yuxi City</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Baoshan City</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Zhaotong City</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Lijiang City</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Pu'er City</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Lincang City</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Cities</td>
<td>2016</td>
<td>2017</td>
<td>2018</td>
<td>2019</td>
<td>2020</td>
<td>synthesis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Due to differences in geographical location, location advantages, political status, resource base, etc., there are differences in the ecological and economic level of various cities in Yunnan Province. Among them, Kunming City remained in the first place and had a significant gap between the score and other prefecture-level cities. Kunming, Qujing City, and Yuxi City basically remained in the top three places, while Baoshan City, Zhaotong City, Lijiang City, Pu'er City, and Lincang City scored lower in comparison. Through the comprehensive score data of 2010-2020, it can be seen that Kunming ranks first among all local-level cities, while Lijiang ranks last in the comprehensive ranking, although Kunming and Lijiang are in an increasing trend as a whole. However, the ecological and economic development of Kunming and Lijiang is quite different.

Therefore, this paper further analyzes the ecological economic development of prefecture-level cities by taking Kunming and Lijiang as examples.

**1) Dynamic scoring analysis of ecological economic development in Kunming**

As the capital city of Yunnan Province, Kunming is located in the center of the economic circle of Southwest China-Southeast Asia hinterland, and has advantages and policies in economic, social and ecological development compared with other prefecture-level cities, so Kunming scored first in the evaluation of the ecological economic development of prefecture-level cities.

<table>
<thead>
<tr>
<th>Year</th>
<th>Economic subsystem</th>
<th>Social subsystems</th>
<th>Ecological subsystems</th>
<th>Comprehensive score</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td>0.000127</td>
<td>0.000206</td>
<td>0.000280</td>
<td>0.000614</td>
</tr>
<tr>
<td>2011</td>
<td>0.000188</td>
<td>0.000258</td>
<td>0.000450</td>
<td>0.000896</td>
</tr>
<tr>
<td>2012</td>
<td>0.000305</td>
<td>0.000323</td>
<td>0.000523</td>
<td>0.001151</td>
</tr>
<tr>
<td>2013</td>
<td>0.00041</td>
<td>0.000249</td>
<td>0.000468</td>
<td>0.001127</td>
</tr>
</tbody>
</table>

Table 4: Dynamic score of ecological economic development in Kunming
It can be seen from Table 4 that the score of Kunming's economic subsystem has increased by nearly 4.75 times in the past ten years, and its economic development level has developed rapidly. The scores of social subsystem and ecological subsystem were relatively stable. Combined with the analysis of the actual situation, in recent years, Kunming's annual GDP has exceeded 600 billion yuan, with an average annual growth rate of 3.0%. The governance of the ecological environment has also been continuously strengthened, the feedback problems of environmental protection inspectors have been dealt with, illegal construction along Dianchi Lake has been rectified, the battle against pollution has achieved a phased victory, the city's ambient air quality has reached the national second-class standard, and the excellent air quality rate has reached 98.65%. While paying attention to the rapid economic and social development of the city, Kunming protects and restores the ecological environment, completing the ecological construction of 55.09 mu of forest and grass, adding 208 hectares of urban green space, and the forest coverage rate reached 52.62%. The above achievements also coincide with the overall growth trend shown by the statistics of each subsystem, but the fluctuation of the score is more obvious and basically in a good state.

(2) Dynamic scoring analysis of ecological economic development in Lijiang City

Lijiang City, as a prefecture-level city integrating the three world heritage sites, world natural heritage and memory of the world heritage, has achieved a huge leap from poverty and backwardness to development and prosperity with the help of unique tourism resource advantages in the past ten years, maintaining rapid and healthy economic and social development in the city, and realizing a historic transformation from agriculture to industry and service industry, and people's quality of life has been continuous improved, the social security system has been continuously improved.

According to the data characteristics of each subsystem in Table 5, the score of Lijiang City in the economic subsystem showed an upward trend, while the score of the social subsystem showed a downward trend, and the score of the ecological subsystem fluctuated significantly, and the growth was not obvious. This coincides with the economic phenomenon that Lijiang City has made full use of natural resources to vigorously develop tourism, doubling its total GDP and exceeding the 50 billion mark, while under the impact of tourism, the ecological environment has suffered certain damage, and the local government has effectively shouldered the responsibility of forest resources protection and development with the "Four Ones". Co-building an ecologically livable environment, which can also explain fluctuations in ecological indicators. However, according to the comparative analysis of the comprehensive score data of Lijiang City and various cities in Yunnan Province, it can be seen that although Lijiang City has
improved its economic, social and ecological development, there is still a big gap with the overall level.

Table 5: Dynamic score of ecological economic development in Lijiang City

<table>
<thead>
<tr>
<th>year</th>
<th>Economic subsystem</th>
<th>Social subsystems</th>
<th>Ecological subsystems</th>
<th>Comprehensive evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td>6.35E-05</td>
<td>0.000204</td>
<td>0.000140</td>
<td>0.000407</td>
</tr>
<tr>
<td>2011</td>
<td>7.47E-05</td>
<td>0.000163</td>
<td>6.57E-05</td>
<td>0.000303</td>
</tr>
<tr>
<td>2012</td>
<td>6.68E-05</td>
<td>0.000204</td>
<td>9.87E-05</td>
<td>0.000369</td>
</tr>
<tr>
<td>2013</td>
<td>7.29E-05</td>
<td>0.000137</td>
<td>0.000144</td>
<td>0.000353</td>
</tr>
<tr>
<td>2014</td>
<td>8.47E-05</td>
<td>8.74E-05</td>
<td>9.34E-05</td>
<td>0.000266</td>
</tr>
<tr>
<td>2015</td>
<td>0.000158</td>
<td>0.000139</td>
<td>8.55E-05</td>
<td>0.000382</td>
</tr>
<tr>
<td>2016</td>
<td>0.000198</td>
<td>0.000109</td>
<td>0.000132</td>
<td>0.000439</td>
</tr>
<tr>
<td>2017</td>
<td>0.000180</td>
<td>0.000106</td>
<td>0.000178</td>
<td>0.000465</td>
</tr>
<tr>
<td>2018</td>
<td>0.000194</td>
<td>0.000104</td>
<td>0.000272</td>
<td>0.000570</td>
</tr>
<tr>
<td>2019</td>
<td>0.000252</td>
<td>9.73E-05</td>
<td>0.000180</td>
<td>0.000529</td>
</tr>
<tr>
<td>2020</td>
<td>0.000258</td>
<td>6.56E-05</td>
<td>0.000176</td>
<td>0.000500</td>
</tr>
</tbody>
</table>

(3) Comparative scoring analysis of ecological economic development in Kunming City and Lijiang City

Based on the data of 2010-2020, it can be seen that the overall development of ecological economy in Kunming occupies a good dominant position among all cities in Yunnan Province, while the overall development of ecological economy in Lijiang City is more worrying, at the end of the score of prefecture-level cities in Yunnan Province, in order to better compare and analyze the ecological economic development level of Kunming City and Lijiang City, this paper draws a comprehensive dynamic score comparison map of Kunming City and Lijiang City.

It can be seen from Figure 1 that the slope of the linear trend of Kunming's comprehensive score is greater than that of Lijiang's comprehensive score, indicating that the growth rate of Kunming's ecological economic development is greater than that of Lijiang's ecological economic development, and through the 2020 data, Kunming's comprehensive score is Lijiang City's comprehensive score is 2.566 times, and the development gap between Kunming City and Lijiang City is large, which shows that the imbalance in the development of various prefecture-level cities in the field of ecological economy in Yunnan Province is more serious.
In view of the reasons for the relatively backward development of Lijiang's ecological economy, this paper summarizes the following three reasons based on the actual situation of the region:

(1) From an economic point of view, for the ecological economy industry, the investment of funds is insufficient and the scale of development is limited. The development of ecological economy requires a lot of investment in ecological and environmental protection infrastructure, the promotion of new technologies, and the restoration of the ecological environment, all of which involve a lot of funds, but due to the shortage of funds, the scale of ecological economic development is limited.

(2) From a social point of view, the foundation of ecological economic development is weak and social participation is not high. The public's awareness of building a resource-saving and environment-friendly society is generally not high, and the sense of participation is not strong, believing that the development of ecological economy takes a long time and high cost compared with traditional development methods, and has no economic benefits and enthusiasm in the short term.

(3) From an ecological point of view, Lijiang's ecological environment is fragile. In recent years, Lijiang's economy has developed rapidly, but in the process of development, it also once relied on energy consumption to develop the economy, and the rough development of high energy consumption and high pollution has had a great impact on the ancient city of Lijiang and its surrounding environment. A series of problems such as the disappearance of animal and plant species. In addition, due to Lijiang's over-reliance on tourism, the eco-tourism environment is deteriorating, artificially putting great pressure on Lijiang's ecological environment. Due to the above reasons, Lijiang's ecological environment has deteriorated, and the economic benefits brought by the ecological environment are low.
4 CONCLUSION

According to the above evaluation and analysis results of the dynamic evaluation and analysis of the ecological economic development of prefecture-level cities in Yunnan Province, the following conclusions can be drawn: (1) The ecological economic development of prefecture-level cities in Yunnan Province has a good trend and a large growth rate. This is inseparable from the economic and ecological policies adopted by Yunnan Province in recent years. (2) The ecological economy of prefecture-level cities in Yunnan Province has uneven development in three dimensions: economic subsystem, social subsystem and ecological subsystem. Among them, it has relatively obvious development advantages in the economy, while the advantages of social development and ecological development are not obvious, and there is a downward trend. (3) The development of prefecture-level cities is uneven, and the scores of prefecture-level cities are quite different. Kunming, Qujing and Yuxi basically maintained their scores in the top three, while Lijiang and Zhaotong lagged behind.

Based on the conclusions reached, this paper makes the following five recommendations:

(1) From the overall results, it can be seen that Yunnan Province needs to closely integrate social and ecological development with high economic growth. On the basis of maintaining absolute high economic growth, Yunnan Province needs to increase the education popularization rate, increase the employment rate around ecological industries, encourage and promote sustainable structural reform of agriculture, industry, service industries and other industries, optimize the energy structure, change the mode of production and consumption, tap the potential of resources, innovatively develop economically developed and ecologically efficient industries, build a culture with reasonable system and social harmony, as well as an environment with a healthy ecology and suitable landscape, and improve the happiness of people's lives.

(2) Yunnan Province can include the ecological economy index in the scope of government assessment, increase the internal transformation momentum, and improve the conversion rate of ecological resources. With the Yunnan provincial government as the main body, encourage the change of the traditional resource-based economic development model, establish a sustainable development technology support system, and form an effective and mature sustainable ecological technology. At the same time, it pays attention to the current situation of unbalanced development of cities at the local level, implements preferential policies and local support for regions, narrows regional gaps, and supports areas with weak ecological economy.

(3) In the process of developing ecological economy, cities at the local level should give full play to their local advantages. Based on the above data, local municipal government organs can make horizontal comparisons, and then go deeper into vertical factors, combine local advantages and characteristics to develop a green ecological economy, and improve the utilization rate of local ecological economy according to local resources.

(4) Improve the plurality of subjects involved in the ecological economy and fully mobilize social resources. Yunnan Province needs to build a diversified investment system and fully mobilize social resources in southwest China. Through administrative and legal means, establish relevant research and development institutions, build a good ecological economic circle, and guide social forces to invest in ecological economic industries. Such as: subsidies for green environmental protection enterprises; Encourage the introduction of professional talents;
(5) Grasp the international situation, respond to the call of the country, and actively develop the ecological economy in the context of carbon neutrality. The state and individuals should maintain a united front and enhance the sense of identity and participation of social groups in building a resource-saving and environment-friendly society.

REFERENCES

Construction of Leader Training System Based on Internet Technology and Big Data Thinking

Jin Li\textsuperscript{1,a}, Qingyu Zhao\textsuperscript{1,b}
lijin@mi.sgcc.com.cn\textsuperscript{a}, zhaqingyu2020@163.com\textsuperscript{b}

Party School of State Grid Corporation of China, No. 12 Mangshanlu, Beijing, China\textsuperscript{1}

Abstract: With the continuous deepening of the reform of state-owned enterprises, the internal and external environment faced by leaders is facing profound changes, which puts forward higher requirements for the leaders and their education and training. The development of Internet and big data technology has brought new optimization direction to the training work. Based on the research on the leaders of energy based state-owned enterprises, this paper analyzes their needs for the design of training system, explores the feasibility of the application of Internet and big data technology in the leader training of state-owned enterprises, and then puts forward the design concept of the leader training system. Through this system, we can provide more scientific and accurate training services for leaders of state-owned enterprises, meet the requirements of enterprises for leader training, and improve the quality of the leader team of state-owned enterprises.

Keywords: Leader Training, State-Owned Enterprise, System Design, Big Data Thinking.

1 INTRODUCTION

Big data is a product of the rapid development of the Internet era, and a new technology derived from the rapid growth of information. Big data thinking is to grasp the internal law or direction of things' development through massive information collection, storage and analysis, so as to coordinate the various elements and relationships between things to solve the difficulties in development\textsuperscript{[6]}. In recent years, the new generation of information technologies, such as big data and cloud computing, have been rapidly developed and applied. The traditional mode of training is facing challenges in many aspects. The use of network technology and network platform for training plays an increasingly important role in building a team of high-quality and professional leaders and promoting the high-quality development of state-owned enterprises\textsuperscript{[1]}. The Regulations on Cadre Education and Training also put forward clear requirements for the use of modern information technology in cadre education and training: make full use of modern information technology, improve the network training system, and establish a compatible, open, shared and standardized network training system for cadres; Improve the informatization level of cadre education, training, teaching and management, and make good use of big data, "Internet +" and other technical means\textsuperscript{[2]}. In order to further improve the training quality and cultivate a high-level leader team, it is necessary to build a leader training system based on big data thinking.

This paper summarizes and refines the feasibility and needs of the leader training system construction by investigating the current situation and needs of leader training in energy based...
state-owned enterprises, and puts forward the concept of leader training system construction based on Internet technology and big data thinking.

2  EMPIRICAL ANALYSIS

We conducted a survey on the training status and needs of leaders of an energy based state-owned enterprise. According to the statistical analysis results of the questionnaire data, this paper summarized and refined the status quo of leader training and the needs for the training system.

2.1  Data Collection

The empirical study reported in this paper focuses on leaders of a state-owned enterprise. The data was collected through an online survey conducted in cooperation with the enterprise. Leaders who participated in the training recently were invited to participate in the survey voluntarily. This survey collected data from different departments and the questionnaire was collected in two times. Through the comparison of the results of the two surveys, it can be considered that the sample questionnaire is representative.

We received 202 valid questionnaires. The majority of respondents (89.11%) were men. In terms of age composition, they were mainly over 46 years old, which is consistent with the basic situation of the main leaders of the enterprise (see Figure 1).

![Figure 1. Basic statistical information of the sample.](image)

2.2  Analysis of Research Results

From the overall evaluation of the training, the survey shows that nearly 30% of the leaders believe that the current training system can meet most of their needs, while the rest of the leaders believe that there is still room for improvement in training (see Figure 2), such as improving the pertinence of courses, enriching teaching forms, etc.
In terms of curriculum resources, 83.76% of the leaders believe that the current curriculum resources are relatively complete and can help them learn better, but some leaders believe that it is necessary to further enrich the curriculum resources. In terms of teaching methods, the respondents reached a high consensus on their preference for case teaching. 91.45% of the leaders prefer case teaching methods, and more than half of the respondents hope to increase the proportion of case discussion (see Figure 3).

In terms of training assistance system, the state-owned enterprise has built an online learning platform. The survey shows that the existing functions of the platform basically meet the needs of most leaders, but the practicality of the course and the richness of information need to be further improved (see Figure 4).
2.3 Analysis of Deficiency

To sum up, the current leaders of state-owned enterprises have the following deficiencies:

- There is a lack of need research, and the training pertinence needs to be improved. Only by conducting training need research can we obtain comprehensive and accurate training need information, determine training content, and design and implement effective training programs. Some units are mere formality in carrying out need research and ignored the job responsibilities and the needs of leaders, which led to the training content and teaching methods not meeting the needs of leaders well, affecting the pertinence and practicality of training.

- Training resources are scattered and cannot be used effectively. The leader training of state-owned enterprises has formed miscellaneous training data. However, the lack of interaction of training data between different units and departments has led to the ineffective interconnection of information and the ineffective use of training resources, which has affected the richness of resources available to leaders.

- The training program is not accurate enough. The formulation of training program is mainly based on the past experience and the unified arrangement of the superior organization department, ignoring the differences in the needs of different posts and leaders, which led to the homogenization of some training and affect the pertinence and effectiveness of the training.

2.4 Analysis of System Requirements

Based on the deficiency of current training, this paper refines the main requirements of leader training system design as follows:

- Combination of dispatch, participation and training. Through the unified training system, users with different roles can achieve the whole process of "superior organization scheduling training", "participating units arranging leaders to participate in training" and "related institutions conducting training".

- Jointly build and share training resources. In order to maximize the overall resource advantage and avoid duplication and inefficiency of resource construction, it is necessary to establish a model of joint construction and sharing of teaching resources. Training data is an important reference for organization departments and participating units to understand leaders, and training resources are an important basis for training program design. The interconnection
of training data and resources across units can help effectively identify resources that meet training needs and output more accurate and effective training programs.

- Precisely output training program. Precision training is an upgrade of traditional training, which helps to avoid formalism and make training more specific and practical [3]. Through the interaction between different user groups, identify the training needs of organizations, units and leaders, match the needs and resources, and output specific training programs to improve the accuracy of training.

3 CONCEPT OF TRAINING SYSTEM CONSTRUCTION

3.1 Design Idea

The design of the training system relies on the Internet information technology to realize the combination of the Internet and the traditional leader training industry, and complete the economic transformation and upgrading by optimizing the production factors, updating the business system, restructuring the business model and other ways.

By using information technology and B/S architecture, deploy servers in the cloud, manage databases and run back-end code; The front-end graphical interface is realized through the applet side, which is convenient for administrators and users to manage and use.

3.2 Overall System Architecture Design

From the perspective of level oriented design, the system design is divided into the following layers: user layer, application layer, platform layer, storage layer, data layer and support layer. The hierarchical structure is shown in Figure 5.

![Figure 5. The overall architecture of the leader training system.](image)

3.2.1 User Layer

It provides a unified portal for three types of user groups include training organization departments, training participating units and training undertaking institutes. Each user group can use this portal as an entrance to carry out training organization, training management, and other related work with the help of the system.
3.2.2 Application Layer

Through the use of various service components provided by the platform layer and the underlying business support system, the training management, learning portfolio management, statistical analysis, comprehensive query, match and program output and other functional applications of this construction are realized.

3.2.3 Platform Layer

As the technical support platform of the whole system, the platform layer needs to provide basic services such as data retrieval system, data sharing and exchange system, data analysis system, data processing system, data and resource integration system, data mining system, and data cloud service system.

3.2.4 Storage Layer

The storage layer adopts flexible and efficient information resource collection methods to realize the centralized collection of original data from different organizations, and realizes the classification and warehousing of original data with the help of data processing methods. Through standardized planning and storage, the original database, trainee database, teacher database, teaching base database and training course database are formed.

3.2.5 Data Layer

Standardize and organize data resources to achieve orderly management of data resources and form a leader training big data directory system, including leader big data, training big data and some unstructured data. Among them, leaders’ big data captures basic files, learning information, training information and other data of leaders; leaders’ training big data captures training resources, training courses, teaching base information and other data; unstructured data covers training related documents, charts, audio and video data.

3.2.6 Support Layer

The support layer includes the standard and specification system, operation and maintenance service system and information security system supporting the stable and safe operation of the system.

3.3 Data Workflow

The leader training of state-owned enterprises will produce a large amount of data, such as historical training information, training needs, training resources, etc. If these data are not processed, counted, classified, analyzed and mined, it is difficult to extract hidden information that can provide support for the development and implementation of training programs, and it will not produce corresponding value.

Effective processing and analysis of data is an important path to achieve accurate matching of resources and needs and improve training quality. With reference to the operation mode of Internet and big data technology, the data workflow of building a leadership training system is shown in Figure 6.
3.3.1 Data Acquisition

Faced with a large amount of training-related data, including leaders' needs, organizational needs, training information and training resources of units at all levels, how to effectively collect, store and analyse these data is a key issue for the construction of the leader training system.

With the help of preliminary research, collect training needs and resources, teachers' information, teaching bases and other data, while crawling historical training data and trainees' files, then enter them into the original database. Subsequently, the detailed data will be received through each database (see Figure 7).

3.3.2 Data Import and Processing

For the training related data from research and crawling history records, there are pictures, tables, documents and data in other forms, which need to be analysed and digitally stored in a structured way while conducting real-time collection. By completing data extraction, data cleaning and conversion, and data loading, the scattered and disparate data collected are integrated together and put into major databases, and the duplicated and redundant information is sifted out.

3.3.3 Data Statistics and Analysis

Through the pre-processing of the data related to the leader training, the original disorganized data can be made convenient for statistics and analysis. The processed training data is further classified and summarized after being filed into each major database.

3.3.4 Data Mining and Application

With the aid of data mining algorithms, a large number of training related data are mined and analyzed to identify the associated data information. The SVM support vector machine algorithm is used to input the different characteristics of each leader, such as age, department, level, training preference, into the established model, and then through parameter optimization,
each person is classified and matched to the most suitable training program for their own, so as to improve the pertinence and accuracy of training.

3.4 User Module Design

The training system provides different functions for different types of users. There are three types of users in this system: training organization departments, training participating units and training undertaking institutes.

3.4.1 Training Organization Department

After logging into the system, the training organization department will enter the organization’s demand for leaders training each year, and delineates the categories of training courses and the minimum percentage of each type of course, and then determine the overall training plan. After the participating units and training providers determine and submit the training program, the program will be approved and fed back (see Figure 8).

![Figure 8. The sketch map of training organization department module.](image)

3.4.2 Training Undertaking Institute

After logging into the system, the training undertaking institute will enter the basic information of teaching bases and integrate the information of available courses, forms and training periods, maximum training scale, and existing teaching facilities and equipment of the teaching bases. These data will be uploaded to provide reference for matching the training needs of participating units and organization departments. If there is a complete training program, it can also be uploaded (see Figure 9).

![Figure 9. The sketch map of training undertaking institute module.](image)

3.4.3 Training Participating Unit

After logging into the system, participating units can enter the basic information and historical training data of leaders. When formulating training program, select the required course in the system and choose a feasible teaching format according to the course, and then select the training teachers and teaching bases that provide the course and format. Subsequently, the training program will be initially output, which can be adjusted according to the actual situation (see Figure 10).
3.4.4 Evaluation

The training evaluation module is open to all users, including the evaluation of training courses, training teachers, training arrangements, overall training effects, trainee performance, etc (see Figure 11). After logging into the system, the training undertaking institute will enter the performance during the training period and the final assessment results of leaders for the training organization departments and participating units to view. For the participating units, they can enter the leader's score for the training program and suggestions for improvement; follow up the leader's work performance after the training, and then provide the unit's score for the training program for the training organization departments and training institutions to view.

3.5 System Deployment and Implementation

This system adopts B/S architecture, and realizes the development and operation of the system in the form of WeChat applet. The background code of the system is deployed in the cloud and written in Python language. The interface is encapsulated using Python's own lightweight web framework Flask. The database uses MySQL to store the data required by the background. The front-end code is developed using WeChat developer tools, and the front and rear end interactions are carried out by directly calling the encapsulated interface in the background, so as to realize the operation of each functional module of the whole system.

At present, the system is on trial in the enterprise investigated in this paper. At this stage, data collection and sorting are the main tasks. More than 400 leaders' files have been entered and training resources such as training courses, teachers and teaching bases are gradually entered. In the future, we will further test the system operation effect and develop more functions such as online training, interaction and message board, so that trainees can have discuss with teachers and other trainees, and then better stimulate their enthusiasm and interest in training.
4 CONCLUSIONS

In recent years, the demand for leader education and training in state-owned enterprises has been increasing with the changing situation, which has brought new challenges to the training organization department. Based on this background, this paper explores the application of Internet and big data technology in the leader training of state-owned enterprises, and proposes the design conception of leader training system combining need research, data analysis, resource matching, program output and training management.

Due to space reasons, this paper only proposes the overall construction conception of the leader training system, which can be tested and improved in future research and practice. In addition, the training of other types of employees can also be included in the operation scope of this training system in future practice.

REFERENCES

A Study on the Application of Digital Technology in Large-scale Events Risk Management

Jiangtian Wang¹, a, Jie Shen¹, b, Xinrui Chen², c, Jiankang Zhang¹, d*

ljiangtian_wang@126.com a, 3227362621@qq.com b, 289995388@qq.com c, *Corresponding author: zhangjk@zisu.edu.cn d

School of Culture and Tourism Zhejiang International Studies University Hangzhou, China¹
School of English Studies Zhejiang International Studies University Hangzhou, China²

Abstract: China advocates high-quality industrial development and industrial digitalization, and digital technology urgently needs to empower the high-quality development of sports service industry. China advocates high-quality industrial development and industrial digitalization, and it is urgent for digital technology to empower the high-quality development of sports service industry. However, there is still no gap in the theoretical and practical research results on the improvement of digital technology technology efficiency in the risk management of large-scale sports events. This paper uses WSR methodology to analyze the application effect of AI in physics, reason and human reason in event risk management, with this mechanism combined with the theoretical framework, specific suggestions for the improvement of scientific and technological efficiency are put forward for the AI technology in the case.

Keywords: Large-Scale Events, Digital Technology, Artificial Intelligence, Event Risk Management.

1 INTRODUCTION

May this year, China gave up the right to host the men's Asian Cup as the organization of the competition was inconsistent with China's concept of putting the first in epidemic prevention. Moreover, in the next ten years, including the Olympic Games, the Women's Asian Cup and the Men's World Cup. China to reduce the occurrence of the number of large events phenomenon is not suddenly, survey nearly 40 years China to undertake the number of large events, ten years ago, the pursuit of quantity has begun to appear fade, this does not mean that the development of large events in our country gradually decline, but our country large events by quantity growth to high quality development gradually transition in the process of profile. We can also see something different from the changes in the development policy of the sports industry. In 2019, The Office of The State Council issued the Opinions on Promoting National Fitness and Sports Consumption and Promoting the High-quality Development of the Sports Industry,¹¹ (No.43), which adjusted the goals and priorities of the development of the sports industry under the new situation. Compared with the "Several Opinions on Accelerating the Development of Sports Industry and Promoting Sports Consumption" (No.46) issued by The State Council in 2014, the current development focus of China's sports industry has shifted from "rapid
development" to "high-quality development". In the specific measures proposed, it is prominent in enriching the objects of policy adjustment and attaching importance to the acceleration of the development of sports service industry. With the advance of the supply-side structural reform, it has shifted from satisfying the demand side to focusing on two dimensions: the guarantee of the supply side and the activation of the demand side. In improving industrial structure, it is proposed to vigorously cultivate service industry, innovate business model and extend the industrial chain. By 2022, the added value of the sports service industry should account for 60% of the added value of the sports industry. Behind the change of the development direction of large-scale events is the repositioning of the functions of large-scale events under the adjustment of the national development strategy. Different political and economic system in China and western developed countries, has long been our country for large events to promote the function of international communication work more attention, due to our country to undertake large event market participation than developed countries, compared to the host city GDP contribution rate can be seen, the role of industry and urban development is relatively limited. With the rise of China's comprehensive national strength and the effective promotion of international communication, the main functions of large-scale events have shifted to vigorously promote the development of cities and sports industry, and large-scale events have moved towards a new stage of high-quality development.

### Table

<table>
<thead>
<tr>
<th>Year</th>
<th>Host Country</th>
<th>Host City</th>
<th>Contribution rate to the GDP of the host city</th>
</tr>
</thead>
<tbody>
<tr>
<td>1988</td>
<td>South Korea</td>
<td>Seoul</td>
<td>1.4%</td>
</tr>
<tr>
<td>1990</td>
<td>Spain</td>
<td>Barcelone</td>
<td>2.9%</td>
</tr>
<tr>
<td>1996</td>
<td>The United States</td>
<td>Atlanta</td>
<td>0.07%</td>
</tr>
<tr>
<td>2000</td>
<td>Australia</td>
<td>Sydney</td>
<td>1.0%</td>
</tr>
<tr>
<td>2004</td>
<td>China</td>
<td>Beijing</td>
<td>0.3%</td>
</tr>
<tr>
<td>2008</td>
<td>The United Kingdom</td>
<td>London</td>
<td>1.1%</td>
</tr>
<tr>
<td>2012</td>
<td>Russia</td>
<td>Sochi</td>
<td>0.83%</td>
</tr>
<tr>
<td>2016</td>
<td>Brasil</td>
<td>Rio de Janeiro</td>
<td>2.1%</td>
</tr>
</tbody>
</table>

**Picture 1.** The contribution rate of the previous Olympic Games to the GDP of the host cities

### 2 LITERATURE REVIEW

Combined with the international research in related fields, it is found that there is still a big gap in the depth of research on how to improve the efficiency of digital technology. In addition, the practical experience at home and abroad inspires us to study from the perspective of supply and demand matching, which is more in line with the current situation of China's economic development and the national development strategy. Zhao Wenting (2022) using the evolution of international sports industry research theory hot spots and practice frontier dynamic, found that the depth of the international sports industry and high and new technology fusion is mainly reflected in areas such as sports competition, and points out that at present our country in this aspect of research and practice is still mainly stay in the concept and form of cross-border accumulation [2]. Dang Ting (2022) analyzed the experience of the digital transformation practice of foreign sports industry, pointed out that the digital transformation changed the core business form of the sports industry from experience-driven to technology-driven, and proposed that China should take the matching between supply and demand as the main direction of the
digital transformation of the sports industry. Li Bo pointed out that China's sports industry has long had a shortage of effective supply and excessive ineffective supply dislocation, and the sports service industry should grasp the current national development of the tertiary industry and the supply side reform to achieve leap-forward development[4]. To sum up, this study tries to explore the mechanism of improving the efficiency of digital technology of large-scale events from the perspective of supply and demand matching in the dimension of large-scale events, in order to enrich the efficiency of diversified digital technology and complex large-scale event management business, and provide reference for the high-quality development of large-scale events.

3 WSR METHODOLOGY ANALYSIS OF AI TO OPTIMIZE THE EFFECT OF EVENT RISK MANAGEMENT

3.1 The WSR Methodological Analysis Framework

WSR methodology [9], namely physical matter and human theory methodology, was jointly proposed by researcher Gu Jifu and Dr. Zhu Zhichang methodology in 1994, and has been listed as "integrated system methodology" by ISSS. The WSR methodology is a system methodology or system thinking method, and it is applicable to complicated systems.

The risk management of large events is a enormous and complicated system, involving complex management objects, covering people, things and things. In recent years, due to inappropriate risk management, numerous unexpected events have caused severe consequences and adverse effects, which have put forward higher requirements for the risk management level of the sports events. However, there is no authoritative assessment system and standard process established in the field of risk management of large-scale events in China, and there are almost no such products in the market. Hikvision AI security products have significant technological efficiency in the event, and have become an efficient risk management tool. Therefore, based on its product characteristics, this paper analyzes the technological effectiveness of AI security in large-scale event risk management with the help of WSR methodology.

3.2 A Brief Description of the AI Operation Mechanism

AI --Artificial Intelligence, In computer science, AI is the intelligence displayed by a machine, known as machine intelligence. Colloquially, the term "artificial intelligence" is used to describe machines that mimic the functions of human "cognition", such as "learning" and "problem-solving". Data-centered machine learning makes data sharing and movement become simple, it regards data as durable basic assets, by compiling a substantial amount of data, the data preprocessing, and then cycle "machine learning modeling-training-evaluation-parameter optimization-training-evaluation-machine learning modeling" process, eventually release application.
3.3 WSR Methodological Analysis of Event Risk Management Efficiency

Hikvision is the official intelligent Internet of Things and big data service sponsor of the 19th Asian Games in Hangzhou in 2022. Asian games organizers signed hikvision, in intelligent command, intelligent security and security, intelligent traffic scheduling, based on several intelligence base power build "panoramic visual, risk prediction, efficient command, agile response" the innovative application of many scenarios, for athletes from various countries and regions, spectators, volunteers, organizers to provide high quality service.

Using AI technology to structure data to create value is the core of Hikvision's products. With video as the core, AI technology is equipped with cutting-edge algorithms to create the overall solution of the Internet of Things and support the linkage between the front and rear end. The huge and rich source data is structured so as to deeply mine the value of data and realize a variety of functions. According to the data based on the AI operation process, the advantages of hikvision AI security series product analysis: first, in the data collection stage, such as "deep eyes" intelligent camera with binocular stereo vision and behavior analysis algorithm, can be extracted in environmental characteristics can extract more rich, more suitable for the characteristics of the parameters, so as to achieve stronger ability to resist environmental interference. Then, in the pre-processing stage, such as the "Superbrain" series NVR, it can extract features with high-order semantics and strong expression ability from the original data, making the accuracy of identifying classified objects higher and effectively improving the value of video. Finally, in the stage of machine deep learning, after training based on enough samples, AI can achieve more accurate target classification identification and autonomous feature recognition, which is especially suitable for abstract and complex analysis of human characteristics and behaviors, and meet the needs of deep-level data value mining. For example, the "Facebook" series of face analysis server, can realize accurate face recognition, 1v1 comparison, identification, control alarm, real-time capture of suspicious targets and other functions. AI security series products have the advantages of comprehensive structured data, deep intelligent algorithm, front and background integration. The data characteristics in the field of risk management of large-scale events are very close to those in the general security field. Therefore, Hikvision's AI security series products can enable the application construction of Asian Games with scientific and technological efficiency in large-scale events and serve the risk management of large-scale events.

From the perspective of WSR system methodology, the risk of sports events can be divided into three categories: physical dimension risk factors, physical dimension risk factors and human dimension risk factors. Next, Hikvision intelligent game watching products analyze the scientific and technological effectiveness that AI security can play in the three risk management dimensions.

3.3.1 Wuli Dimension

Wuli in WSR methodology refers to the objective existence that people face in the problem processing process of a certain system project, which is the sum total of the laws of material movement. The threat factors at the wuli level of large-scale events mainly include environmental variation risks, venue and facilities risks, and event protection risks. In the Asian games, AI security can be based on the cloud platform, implementation and Hangzhou use city, transportation wisdom, wisdom, wisdom medical services such as data sharing and circulation,
to the environment of urban intelligent monitoring power events of natural disasters timely warning and continuous tracking outbreak hierarchical control, implementation of environmental variation risk timely warning and processing. Hikvision's "Facebook" series of face analysis server, can realize accurate face identification, 1v1 comparison, identification, control alarm, real-time capture of suspicious targets and other functions. AI security with video stereo vision combined depth intelligent algorithm layout wisdom, can control the location construction quality, rich venues wisdom function, rapid scheduling facilities, etc., for security and location epidemic prevention [9] and control, AI security in addition to basic face recognition, also can intelligent analysis across the cordon, wandering, running detection, the number of abnormal detection abnormal detection, spacing detection, strenuous exercise, leave detection, ground detection, stranded nine behavior, so as to realize the venue facilities risk and event security risk prediction, real-time control or avoid.

3.3.2 Shili Dimension

The principle of WSR methodology refers to the mechanism of objective existence and its intervention in the process of handling the problems of a certain system project. The risk factors at the wuli level of large-scale events mainly include event operation risk, organization and management risk, and information dissemination risk. AI + security innovative camera technology and cutting-edge algorithm can adapt to various scenarios, meet the needs of high-speed camera, diving, ball and other different events, zoom, precision, trajectory prediction and other functions, assist judges to make a more fair and more credible judgment, and reduce the risk of event operation. Hikvision's "Superbrain" series NVR can extract features with high-order semantics and strong expression ability from the original data, which makes the accuracy of identifying classified objects higher and effectively improves the value of video. AI security multidimensional information in the form of video, can greatly reduce the risk of distortion in the process of information transmission, blocked, and in the background algorithm for intelligent analysis, for decision makers to collect more comprehensive, objective, structured information, in various departments, complex process of large event management process, auxiliary science, efficient decisions, and based on cloud platform to realize real-time information sharing and rapid transmission, improve the efficiency of upload and department collaboration, realize timely response to emergencies and rapid and scientific processing, avoid the risk of information dissemination, and organization management risk.

3.3.3 Renli Dimension

Renli principle in WSR methodology relates to the interrelationship and the changing process of all the people involved in the problem processing process of a certain system project. The risk factors at the wuli level of large-scale events mainly include personnel relationship risk, financial management risk and legislative risk. On the basis of video surveillance, AI security can analyze human behavior and face expression, identify abnormal behaviors or crowds, and predict unexpected crowd conflicts or riots in advance. Video as the core of the AI security through the camera as the intelligent terminal of the Internet of things, build event scene digital twin, on the basis of insight, control, optimization of management and creation, timely meet the different national government personnel, coaches, athletes and other event participants demand, provide more tailored, humanized service, at the same time reduce the relationship and financial
risk. AI security linkage real-time camera and background intelligent analysis and storage, for tournament interrogation institutions provide intuitive and accurate evidence to assist judgment, to support the police in different scenes of face, behavior, traceability, will involve legal, moral, ethics, event cases as far as possible to reduce the influence of the impartiality, reduce the risk of event legislation.

4 MECHANISM REFINING

Event management and digital technology are characterized by high research and development risk, high application cost and long preparation cycle. The current efficiency improvement of shallow and decentralized digital technology will cause two problems: first, the new technology and the need to continuously run in with the old business process, which limits the full play of efficiency; second, the ecological formation of technological innovation and service upgrading, the lack of endogenous power, the sustainability of efficiency improvement is limited, and the process of industrial digitalization will be blocked. Even if the breadth of digital technology application coverage increase, various digital application scenarios, level between no systematic organic combination, then digital enough deep, efficiency can only be a form of "smart", does not achieve the true sense of "wisdom", events of high quality development may be easy to encounter bottlenecks. Reference to other industry digital experience, digital technology innovation is the industry the original management, marketing, system, process, concept of innovation, rather than the digital technology only as a new tool to use, single, the surface of the innovation cannot form a positive cycle, promote the industry sustainable development, efficiency promotion is also very superficial, short.

4.1 Mechanism Advantage

The internal and external double-ring mechanism is characterized by the dual subjectivity of the elements and the mobility of the ring structure. Its advantage is that the double subjectivity to comb the main body of supply and demand, has been clear about the fundamental demand of event management link, and the event management link execution process to distinguish as a supply factor, no longer just digital technology one-way to serve the existing execution process, but emphasize two supply factors dynamically mutual adjustment, jointly form supply products, to meet the demand of event management link at all levels. This can not only stimulate the improvement of event management efficiency, but also enable digital technology to give full play to different advantages at different levels. In the process of continuous adjustment and matching of digital technology and event management, the degree of digitalization of large-scale events will be deepened. The liquidity of the circular structure comprehensively considers the internal and external factors affecting the supply and demand, and enriches the dimension of adjusting the matching between supply and demand. The value generated is constantly accumulated in the closed loop to form the driving force to promote product innovation, and supply and demand tend to match, making sustainable development possible.

For digital technology, multi-level demand expands a great space for technology play. Accurate demand and timely application feedback provide direction guidance for continuous innovation, which is conducive to the improvement of scientific and technological efficiency at the
application level. Continuous funding sources and rich application scenarios can promote the technical efficiency of its technology.

4.2 The Power of the Mechanism in Operation

4.2.1 Elements and Factor Structure

Endogenous motivation is the value realization of event management and digital technology. In the process of pursuing supply and demand matching, the efficiency of competition management has been continuously improved, digital technology products have been constantly optimized, playing a role in large-scale events to realize technical value and increase the value in the market, thus promoting the continuous innovation of digital technology and improving the efficiency of science and technology. The external push is investment by markets, the government and society. External tension is the consumption of the market and the expansion of the downstream application of the industrial chain. For example, after the use of the venue, some technology transfer during the competition is applied to leisure, education, education, health care, etc. Large-scale sports events are located in the upper and upper reaches of the sports industry chain. High-quality sports event products can accumulate value along the industry chain and drive the development of the downstream and even the whole industry.

5 AI STRATEGY OF IMPROVING TECHNOLOGY EFFICIENCY IN LARGE-SCALE EVENT RISK MANAGEMENT

As for the improvement of technology efficiency of digital technology in large-scale events, this paper believes that from the perspective of supply and demand matching, digital technology optimizes supply to guide the demand of large-scale competition management, and to expand the breadth and depth of digital technology efficiency in the application of different levels with hierarchical demand. Specifically, through with large event management link execution process mutual adjustment, respectively meet the demand of all levels, improve the
supply and demand of matching degree and on the whole application level to improve the efficiency of digital technology of digital technology, to a certain extent, promote technological innovation, improve the technical level of science and technology efficiency.

5.1 AI Security Based on WSR Methodology in the Risk Management of Large-Scale Events

5.1.1 Wuli Level

When using AI security to layout the smart venues, on the basis of ensuring the operation function of large-scale events, it can be combined with the post-game utilization planning of the venues in advance, and flexible and diverse functions can be designed with digital technology to improve the economic and social benefits of the smart venues. Such as the application of digital technology to the mass sports activities, let mass sports events can also have event level AI referee and AI photography, AI coach and emergency medical reduce the risk of injury in sports, sports risk management [6] service mass sports risk management, with wisdom power health, to event level risk management to ensure the safety of mass sports. The transfer and application of the digital media technology of the events, the three-dimensional live broadcast of large-scale evening parties and concerts, the flow of people and risk management of immersive performances, etc., will meet the needs of the venues to transform into sports and sports centers after the competition, explore the diversified ways of profit of sports and sports space, and reduce the operation risk of the venues. And suggest wisdom venues through the cloud platform access wisdom city system, as the city of "wisdom assets" flexible scheduling, through the AI security deep learning all kinds of natural disasters, outbreaks and other force majeure risk plan, in order to quickly equipment become "square" wisdom, timely service urban governance and urban risk management. AI security enables smart venues, to precipitate and expand the technological efficiency of the digital technology in the event after the competition, and to return it to the people, and to magnify the social benefits of the digital technology of the event.

5.1.2 Shili Level

![The WSR methodology workflow](image)

*Picture 3. The WSR methodology workflow* [5]
Making comparison, the workflow of the WSR methodology is similar to the operation principle of the AI: collect information (data) - information (data) structured (modeling) - implementation scheme (machine learning) and test regulation (regulation parameters) cycle-implementing the final protocol (release application). Therefore, it is suggested to take AI security as the digital intelligence base, build the AI event risk management system based on WSR methodology, and design the risk management operation mechanism of large-scale events based on the workflow of WSR methodology.

The AI event risk management system needs to be characterized by fast response speed, high degree of information structure, strong department collaboration ability and growth of the system, and it also needs to improve the management efficiency of large-scale events and optimize them in specific ways and work processes. Before the competition, digital technology was applied to comprehensively collect the requirements of the event, analyze and plan the event with physical-reason-human structure, use AI to continuously evaluate and adjust the planning scheme, and finally select satisfactory plans, so as to minimize the consumption of manpower, material resources and time cost caused by inappropriate decisions. During the competition, the intelligent command based on the Internet of Things and cloud platform helps high-level decision makers to achieve three-dimensional perspective of information, vertical important decisions can be transmitted quickly, and horizontal personnel of various departments can share information in real time, flexibly dispatch resources, and improve collaboration effectiveness. At the same time, the effect of the decision execution is monitored in real time, and the instructions are adjusted quickly through feedback to achieve the optimal decision of physical-matter-human dimension. In addition, AI can be used to restore the process of decision paths and organizational action. Through machine learning, a management effectiveness evaluation system can be established to identify the links and behaviors that need to be optimized, so that the event managers can optimize the workflow of event management and organizational cooperation from an objective perspective, and improve the management effectiveness.

5.1.3 Renli level

Large-scale event level high, project, there are many official organizations and authorities to participate in, the international attention and influence is very high, so the event is not only a
high-end sports competitive platform \cite{7}, but also an important window of spirit and culture communication, event to provide services and event management is the participants to convey the spirit of the most intuitive experience. Through AI deep learning of a large number of videos of the Paralympic Games and disabled athletes, the live habits and behavior characteristics of special spectators, and integrating the humanized facilities \cite{8} for the smart venue, adjust the venue facilities based on video monitoring and face recognition, provide real-time following voice navigation for visually impaired people, monitor and clean up the venue channels, and ensure the safety and free activities of people with mobility disorders. AI event management through the application of digital technology, in the event planning to sustainable development, reusable principle design venue function management mode, extend in time, space on the economic benefits and social benefits, accurate events in the operation process of the regulation, save unnecessary manpower, material waste, after the game to evaluate event management, analysis and summary, formation methodology rich event management case and mode, continuous improvement provide reference for the next event. Moreover, the event management with video as the carrier and digital technology as the core can extend the management chain in the industrial chain, and interact with multiple roles such as audiences, sponsors and clubs by intervening in the media, marketing and other segments of the event. The peaceful call, the sports spirit of the event, as well as the humanized intelligence, environmental protection and sustainable advocacy of the event, so as to realize the cultural efficiency improvement of the event with the technological efficiency of digital technology.

6 RESEARCH CONCLUSION

In general, this paper combed the domestic and foreign to explore large event digital technology efficiency promotion strategy of the literature, determine the supply and demand matching for research perspective, through the analysis of case, combing and innovation of digital technology and competition management of supply and demand matching mechanism and digital technology science and technology efficiency promotion strategy, and use the mechanism combined with case the AI efficiency of science and technology improve specific Suggestions.

- AI security helps with the use of smart venues after the games to improve the social benefits of large-scale events.
- Use AI risk management to promote event management reform, and improve organizational management efficiency with WSR workflow.
- AI enables to optimize the service of special groups of the event, and improves the cultural benefits of the event with video as the carrier.

Due to the limited time and ability, there are still many deficiencies in this paper, such as the need for first-hand data and data for empirical testing to correct the mechanism. In the context of high-quality development, it is hoped that with the joint efforts of scholars, digital technology can be deeply combined with large-scale sports events, to promote the rapid digital development of the sports industry, and lead the development of the city.
REFERENCES


The Development of Digital Technology Efficiency in the Communication of Large-Scale Events

Jie Shen\textsuperscript{1,a}, Jiangtian Wang\textsuperscript{1,b}, Jiankang Zhang\textsuperscript{1,c*}

3227362621@qq.com\textsuperscript{a}, jiangtian_wang@126.com\textsuperscript{b}, \textsuperscript{c}*Corresponding author: zhangjk@zisu.edu.cn\textsuperscript{c}

School of Culture and Tourism, Zhejiang International Studies University, Hangzhou, China\textsuperscript{1}

Abstract: Under the wave of digital technology revolution, smart sports events are leading the development of Chinese and international sports events in a new direction. However, there is still a gap in the theoretical and practical research achievements on the improvement of digital technology in the communication of large-scale competitions in China. Article uses stack poly-spoke model to analyze the application effect of AR in symbol, emotion and situation in the event communication. And Starting from the operation mechanism of the model, it provides a strategy for improving the application efficiency of digital technology in the event communication.

Keywords: Large-Scale Events, Digital Technology, AR, Event Communication.

1 INTRODUCTION

Under the context of digital economy as the main economic form, digital technology has been applied in large-scale events more and more widely, and has become one of the necessary means to ensure the normal operation and efficient dissemination of events. The Hangzhou Asian Games has also included "intelligence" in the concept of the Asian Games. In October 2021, the General Administration of Sport of China issued the 14th Five-Year Plan for Sports Development, which proposed to "support the innovative application of new technologies such as big data, blockchain, Internet of Things, cloud computing and artificial intelligence in the field of sports". In order to meet the new demands proposed by the transformation of large-scale events, digital technology needs to transform scientific and technological tools into scientific and technological efficiency and become a new engine for the high-quality development of large-scale events. On the basis of improving the management efficiency of large-scale events, it can be transferred and applied to the sports industry and energy development in cities. Efficiency is different from efficiency, which is not only reflected in the size and level of effect and benefit, but also more importantly reflects the size and strength of the effect and benefit ability, improves efficiency to grab the optimization of elements, and the mobility and self-growth of efficiency put more emphasis on people's subjective initiative. Large event digital technology application has the development difficulty, high cost, long
recovery cycle, and large event preparation cycle is long, short duration, the fault tolerance rate is low, subject to the digital technology and large events basic features, the external advantage of digital technology limited, and digital technology and large events between professional information permeability is low, only rely on technological innovation for large events of digital technology efficiency is not significant. Therefore, how to improve the efficiency of digital technology in large-scale events is an important task for the high-quality development of large-scale events, and also the need for the high-quality development of sports industry and competition cities, which has become a topic of significance and value at the practical and academic level.

2 LITERATURE REVIEW

At present, there are few research documents on the improvement of digital technology efficiency for large-scale events in China, but there are many documents on the digitalization of sports industry, sports service industry and sports communication with large-scale events as the carrier, which can be connected with them. Overview of related literature, for efficiency promotion, cut into the perspective of focus on the sports industry development and events, based on the existing product case analysis put forward targeted improvement strategy research way, improve efficiency strategy is mostly optimized application environment or combined with technological innovation, to expand the application scenario of digital technology.

Shen Keyin proposed that the digital development of sports service industry should strengthen the top-level design and improve the development environment in the government dimension, gather factor resources in the industrial dimension, and promote the supply-side structural reform [1]. Chai Wangjun sorted out the function mechanism of specific digital technology to enable the high-quality development of sports industry, and analyzed the application dimensions [4] of various digital technologies in different sports service segmentation industries from the characteristics of digital technology perspective. Wang Xiangfei, combed the specific application of VR in large sports communication, analysis and put forward in the transmission of image production, social transmission, equipment compatibility problems, and based on this, from the communication technology, industry standards, talent training, virtual reality technology in large sports events communication application of optimization strategy [5]. In addition, some studies also explore the application of digital technology in large-scale events from the technical dimension, such as the application of AI intelligent technology [6] and multi-dimensional perception system [7] in the security link of large-scale events.

3 HOW DIGITAL TECHNOLOGY PLAYS THE EFFECTIVENESS OF TECHNOLOGY

Beijing New Olympic Group, which takes video product development and production and system integration as its core business and takes commercial real estate operation as the value-added mode, and provides live broadcast technical services for China's 2022 Beijing Winter Olympic Games. New Otter AR system —— Graphite Ultra HD online virtual packaging system provides the text broadcast for the whole HD of the event through uHD virtual
implantation technology and uHD large screen. In terms of uHD virtual implantation technology, it mainly uses three core technologies: sensing and positioning, virtual layered rendering and synthesis, and GPU color keys. Sensing positioning uses image recognition technology to locate the position and posture of the camera in a virtual environment, Ensuring the accurate positioning of the camera in the environment while also ensuring the accurate capture of real-time images of the field, To meet the audience's demand to watch the wonderful moments of the athletes in the event; Virtual layered rendering synthesis technology through the combination of virtual broadcast, virtual implantation scene and packaging layer to achieve rich broadcast effects, It is conducive to the audience's pursuit of real-time broadcast and camera review of the game score; The GPU color key technique involves the processing of scene brightness and shadows, Make the real people in the virtual scene more "sense of landing", Enrich the audience's visual experience of the saturation picture of the game. In the audience survey the influencing factors of the experience, the audience to the platform quality definition has high requirements, new otart, ultra-high definition screen not only provide 5G + 4K / 8K images, also in data input processing, it docking internal and external data sources for project control, make the event forecast, the game table, athletes real-time situation, the team formation data visualization, visualization.

3.1 The Aggregation-Convergence Model to Analyze the Communication Effect of AR Optimization

3.1.1 The Analysis Framework of the Stacked Poly-Convergence Model was Established

The stack poly-convergence model is a model of analyzes traditional media information by using the way of information transmission and the evolution rules from a decentralized perspective. The stack-convergence model distributes the multimodal aggregated information to radiate the media triggers to the audience through three dimensions of symbol, emotion and situation, and the audience produces different information behaviors at different acceptance nodes. The superposition-divergence model is divided into two processes: aggregation and divergence. The operation mode of the two processes is shown in the figure below. Introduce the model of AR, VR technology watching products to explain, fold poly more focused on AR technology interaction, can through intelligent products based on the reality environment of virtual objects, information display to pass to the change of the audience, and spoke more focused on AR and VR technology to the audience virtual information interactive entertainment and panoramic view of the virtual world.
Through the manifestation principle of various colors in the traditional optical mode, especially through the difference of red, green and blue, the signal composition of the new media transmission is examined from the perspective of overlapping and radiation, so as to study the fusion law and transmission effect between the signal components of the three dimensions such as symbol, emotion and situation.

3.1.2 AR and VR Technology Operation Principle and the Two

AR ——Augmented Reality, AR generally refers to augmented reality technology. AR information trigger is AR products in information to enhance the trigger mechanism, namely in the visual information conveyed by the real object increased cannot touch the virtual information, through AR product stack poly reality and virtual fusion information, radiation out a more complete useful information. Under the stack-convergence model, the information trigger of AR is the trigger mechanism of AR products when the information is enhanced, that is, the untouchable virtual information is added to the visual information conveyed by real objects, and through the overlapping reality and virtual fusion of AR products, more complete and useful information is distributed.

VR ——Virtual Reality, VR generally refers to virtual reality technology. VR technology is the use of computer simulation technology to create a virtual three-dimensional world, to provide users with visual, auditory and tactile sensory simulation, through a variety of paths to text, pictures, sound contained in the information conveyed out.

The difference between the two is that AR technology emphasizes the interaction of information, on the basis of the real world, add or remove virtual objects generated by computer or information that can interact in real time; VR technology emphasizes panoramic interaction, a complete virtual three-dimensional world created by computer simulation technology similar to the real world but isolated from the real world, and three-dimensional immersion brought by the virtual environment by mobilizing human senses.
3.2 A Superposition and Divergence Model Analysis of Event Communication Efficiencies

This paper analyzes the overall efficiency, operation mode and information dissemination mode of virtual intelligent competition watching products based on innovative technology from the three levels of symbol, emotion and situation.

3.2.1 Symbolic Level

Currently, people are already stuck in the symbolic landscape created by social media. The meaning of medium existence is reality, namely symbolic life. Since the advent of smartphones, the world has changed. The symbolic world built by social media on mobile terminals completely binds people together. In the foreseeable future, the emergence of wearable devices, virtual reality, and augmented reality will make people more deeply involved in the media survival.

At the symbolic level, AR enriches the media of event communication. During the watching process, the spectators, as the recipient, can combine the indirect media of AR and VR terminals to approach the real perception of the human body, that is, the direct media. Indirect media is that the audience is not on the competition scene. At this time, the media is more inclined to watch the event broadcast products, such as mobile phones, tablets, computers and large-screen TV tools. For a series of watching products, whether broadcast tools or live watching wear can enter the virtual world and athletes interactive VR glasses, can let the crowd access to the event process, athletes, and focus, cause communication between domestic and international discussion, bring language speech and thinking together color.

AR / VR also provides an effective path to grasp the transmission effect. In the dimension of symbolic information, the amount of information propagation is calculated by the similarity of the subject vector between the information, and the evolution of media information propagation is viewed from the principle of light overlapping, so the amount of comments is the symbolic dimension in the propagation of media information. The audience forms multiple communication sub-groups on different media platforms. For the behavior data of their comments, the communicators of the event can screen the data and screen out the comments with virtual technology products such as "AR and VR", and screen the audience's product evaluation or suggestions again, and make improvements according to the relevant useful information.

3.2.2 Emotional Level

The effective application of AR / VR intelligent game-watching products takes social diversity and enhances people's emotional experience.

AR / VR promotes personal emotional expression and group emotional transmission among spectators. The psychological elements of the users of digital technology viewing products predominantly include individual psychological and social psychological elements such as cognition, emotion, personality and interpersonal communication. In the social media environment, emotion is the fundamental power of sports topic event heat, sports events for the audience emotional transfer is to form the root cause of large-scale sports mobilization, the
audience emotional reaction to the events of social cultural connotation, in the form of emotional resonance and discourse together to participate in public opinion. From the perspective of the prototype theory, emotional transmission and prototype contains the group sharing the emotion and meaning of mutual overlap, driving the generation and development of public opinion, and all kinds of AR individual virtual image and expression helps to make more intuitive reveal watching emotion.

The AR + social scene has been widely used in sports events. The application of AR + social interaction introduces people's emotion of watching games into the research of virtual space. The psychological elements of users of digital technology watching games mainly include individual psychological and social psychological elements such as cognition, emotion, personality and interpersonal communication. In the social media environment of events, emotion is the fundamental driving force of sports topic events, and the mobilization of sports events to audiences 'emotions is the fundamental reason for the formation of large-scale sports mobilization. The audiences' emotional reactions to events receive the marketing of social culture, and participates in public opinion in the form of emotional resonance and discourse coordination. From the perspective of the prototype theory, emotional communication and prototype contains the group sharing the emotion and meaning of each overlap, driving the generation and development of public opinion, and AR + social scene application to share the emotion and meaning deepening amplification, all kinds of AR personal virtual image and expression made more intuitive show watching emotion.

3.2.3 The Situation Level

The intelligent XR viewing products provide the audience with an interactive time and airspace based on emotion. The "boundary" of the situation has not only the wuli definition domain such as "boundary and boundary", but also the time and airspace in the wuli sense of "local and regional". The intelligent game-viewing products break through the boundary in terms of the situation and provide the network virtual space, so as to create a platform for the audience to communicate emotionally through the virtual world provided by XR tools. In terms of e-sports and social interaction, VR video games provide users with virtual real scene dialogue, shooting, fighting, decryption, installation art and other scenes to create sensory stimulation. In terms of film and live broadcasting, AR special effects are the exclusive experience of non-site users. Virtual reality / enhanced expression forms portray the scene according to AR / MR special effects application to present the visual aesthetic arena for the audience.

![Figure 2. Context and situational information diagram](image)

3.2.4 Symbol, Emotion, Situation Level Progressive Analysis

This paper holds that the symbol, emotion and situation form a progressive relationship in the communication and interaction of the event. Through the design and development for watching the event and the audience, the emotion transmits the symbol, and arouses the audience emotion.
to embed the emotional color for the interaction between people and the environment; the situation increases the symbol and emotional additive effect, thus providing a communication platform for the interaction between people and the environment. The progressive combination of the three provides the fundamental ideas for the application of digital technology in large-scale competitions.

4 MECHANISM REFINING

Based on the analysis of the above, it can be found that in the practical application, the core requirements of the event management link and the principle of digital technology are not easy to change, but the management process and the application form of digital technology generated to meet the core requirements are not fixed. The two are the relationship of mutual promotion and coordination. Therefore, it is necessary to combine the principle of digital technology with the fundamental goal of specific business and the underlying logic, and redesign the business process, which is more conducive to realizing the overall maximization of management efficiency and science and technology efficiency, so as to improve the efficiency of science and technology at the application level. In view of the complex process of the interaction between the two, on the basis of sorting out the matching process of supply and demand, a new mechanism for improving the efficiency of the efficiency of digital technology in large-scale events from the perspective of supply and demand matching is explored and proposed.

4.1 Two Main Contents of Supply and Demand Matching

4.1.1 Static Matching

First of all, combining the characteristics of a certain link of the event management, mining its most fundamental needs, and matching digital technology that can play the most efficient role in principle. For instance, the prominent feature of the event communication link is the variety of communication media, and its fundamental demand is to alleviate the information asymmetry between the subjects, while the technical advantages of AR augmented reality and VR virtual reality can improve the rate of information restoration and the richness of communication media, and highly match the needs of the event communication link.

4.1.2 Dynamic Matching

Digital technology enriched the event management link the realization of the fundamental goal path, should be event management link business process and digital technology as dynamic change supply factors, in order to meet the fundamental demand of event management link as the guidance, and at the same time adjust the event management link specific execution process and the application of digital technology, improve the matching degree of supply and demand, in the dynamic adjustment, find management efficiency and technology efficiency of overall maximization.
4.2 The Structure of the Supply and Demand Dynamic Matching Mechanism

Elements: digital technology and event management links (such as event communication), specific implementation process of digital technology products and event management links, large-scale event service products, funds, markets (including digital technology products market and large-scale event service products market), government and society.

Element structure: The relationship between the elements can be described as the double-ring structure presented in the figure below. In the interior ring, event management link and digital technology on the basis of static matching, through innovative digital technology products and redesign or optimize event management link specific execution process, common iteration to mutual assignment, in the process of sustainable dynamic matching, high quality large event service products as output, meet the market, government and society proposed different demands for large events, to achieve value or even value-added. As feedback, the market, government and society in the outer ring inject capital into the interior ring, accelerate the capital flow of the inner ring, and promote the upgrading and iteration of the inner ring elements. In this way to stimulate the vitality of the interior ring and the internal elements of the inner ring, promote the circulation of the inner and outer ring, and form a sustainable operation mechanism.

Figure 3. Structure diagram of the double-ring elements

Figure 4. Structure diagram of the double-ring elements
For digital technology, multi-level demand expands the space for technology play, etc. Accurate demand and timely application feedback provide direction guidance for continuous innovation, which is conducive to the improvement of scientific and technological efficiency at the application level. Continuous funding sources and rich application scenarios can promote the technical efficiency of its technology.

5 AR + VR STRATEGY OF IMPROVING TECHNOLOGY EFFICIENCY IN LARGE-SCALE EVENT COMMUNICATION

5.1 Symbol Level

With the continuous progress of digital technology, people's requirements for competition commentary are getting higher and higher. As the bridge and link of sports events, commentary has the functions of connecting the preceding and the next, explanation and technical analysis in the dissemination of sports events, so that people can understand sports while popularizing sports knowledge, improve people's interest in watching, guide the masses to participate, so that the masses can fully feel the unique charm of sports events. In the studio, AR technology can not only change the environment, but also utilize the virtual screen to provide the competition field, athletes, action analysis and other information on the commentary of the commentator, so that the audience can understand more comprehensive information about the event on the spot in real time under professional commentary.

5.2 Scenario Level

The research report on the users of the 2021 Tokyo Olympic Games has divided them into four categories, namely, the national team iron fans, the vertical fans, the competition spectators, and the family companions. National iron fans attach great importance to the competition, Positive or negative mood differences generated by the title or defeat of the favorite national players, From the analysis of the stacked-radiation model, These differences give an emotional color to the event; Vertical lovers value the process of the event and the highlights of the athletes, belongs to the situational level of the model, This group has high spatial and temporal requirements for the situation, The process of AR augmented reality technology and VR virtual space technology, To meet people's demand for camera-capture moments; The crowd are more inclined to the results of the event and off-site news, beginning from the symbolic dimension in the model, Using AR recognition technology applied in the event to provide better information on the latest trends of the event. Let the audience receive more comprehensive information; Family companions attach importance to children's patriotic education, VR glasses can be used to guide their children to the virtual environment that simulates China's early participation in the World Games to experience the difficulties of people at that time. And to conduct sports knowledge and popular science for their children.
5.3 Emotional Level

New Zealand Internet software Dropit has developed a 60-second auction on the stadium score card to enhance the interaction between the brand and fans. During the game, AR + purchase app can be set up to buy interactive items. Viewers can bid for the individual sports time of the winners, or buy the signature photos and commemorative T-shirts of their favorite players. 3d scanning technology is applied to 3D data the surrounding athletes, the same clothes, wristbands, hair accessories and other items, and set the small purchase logo at the time of the event for the audience to click and purchase. Interaction is an important part of immersive game watching. We have to mention the virtual technology here. According to the survey results, more than half of the Internet users are willing to try the technology, while 30% of the netizens will wait and see to become potential users of the technology. For people who can't be there, virtual worlds provide an excellent interactive platform.

With the in-depth development and extensive application of VR technology, it is estimated that by 2025, the total market volume of VR content in China will exceed 83.27 billion. In the next five years, the world VR content field investment growth rate is the fastest enterprise training, and then the game industry, but the sports industry is not on the list, by 2020, the world of VR / AR project investment and financing amount is about two hundred million yuan, 2019 rose 14 percent: each link from the world of VR / AR project investment and financing, hardware and application link is the hot link of investment and financing, this shows that the application of VR in the event has a great development space.

In order to further improve the application of digital technology in the event and enhance the interactivity of people watching the game, applicable digital research departments can use virtual technology to provide each participant and the audience logged on the official platform with DIY own image platform, the audience can use VR glasses to watch the game in the virtual world stream interaction.

VR technology has also brought benefits to people with disabilities. VR helmets developed by an innovative company in London can help people with visual impairment to get their eyes back to regular levels, and enable people with visual impairment to watch wonderful sports games. Virtual images can also create different experiences to people with mobility difficulties. Deaf and mute people can also connect the Asian Games site through AR glasses and use double screen and one screen to present high-definition scene and subtitle accurate translation screen for game interaction, enriching the game watching experience of impaired people in an all-round way. As a platform, these functions can be collected according to the needs of the audience to facilitate people's use and switch of various functions.

6 RESEARCH CONCLUSION

As for the improvement of the efficiency of digital technology in large-scale events, this paper believes that from the perspective of supply and demand matching, the demand to guide the supply of large-scale events, and expand the breadth and depth of digital technology efficiency in the application of different levels with hierarchical demand. Specifically, through with large event management link execution process mutual adjustment, respectively meet the demand of event management link at all levels, improve the supply and demand of matching degree and on
the whole application level to improve the efficiency of digital technology of digital technology, to a certain extent, promote technological innovation, improve the technical level of science and technology efficiency.

Today's large sports events more interactive expanding, melting media era has come, watching the crowd before the event, in the late interactive entertainment experience requirement is becoming higher and higher, according to the results of the survey, 83.24% of the audience will share each other watching feelings, video platform and social platform to become the masses, sharing, evaluate the main channel of empathy. The main forms of mass game watching are live broadcast, short video, replay, full audience viewing, round broadcast, etc. The linkage of large screen and large TV screen on the mobile terminal have become the mainstream of game watching tools. At the same time, the mobile terminal also provides a social platform for the audience to interact and discuss the event. During the discussion, the audience tends to discuss the performance of the national team, opening and closing ceremonies and performances, the referee's decisions and disputes, and the prediction of competition results, etc. In the preference of these discussions, the public believe that the main factors affecting the interactive experience on the platform are: content richness, picture quality clarity, professional interpretation, viewing fluency, function richness, etc. Therefore, to enrich the platform content, invite more professional commentators, make the platform picture quality clearer, and improve the relevant functions of the platform to improve the user experience.

In the process of watching the game, the functions they most want to add are real-time broadcast, VR panoramic viewing, intelligent portrait recognition, screen recording function and interactive entertainment, which shows the increasing demand for independent operation and freedom. As the platform side, these functions can be concentrated on one platform to facilitate the use and switching of various functions.

Therefore, from the following three points for digital technology in the dissemination of large-scale events recommendations.

- Text, pictures, sound and video are integrated, professional commentary to deliver the wonderful information of the event.
- Subdivide the watching groups to provide personalized customized services for groups with different needs.
- Interaction promotes the game watching experience, and uses the virtual technology to create a diversified digital interactive platform.

In general, this paper has the following three innovative points: the first is the new perspective, explore the large event digital technology technology efficiency improvement strategy from the perspective of supply and demand matching; the second is the new mechanism: deconstruct and reconstruction of the digital technology and event management supply and demand matching mechanism; the third is the new suggestion: using the innovative supply and demand matching mechanism to with the specific suggestions of AR technology efficiency improvement.

This paper is the phased achievement of the follow-up development and utilization of Hangzhou Asian Games venue of "Xin Miao Talent Plan" of the Zhejiang Province.
REFERENCES

Research on the Association of Digital Media Art and Cultural and Creative Industry

ShuYao He*, JingHong Chen
*160507207@stu.cuz.edu.cn
Hanyang University, Ansan, Gyeonggi-do, Korea

Abstract: With the rapid development of computer today, digital media art has also developed accordingly. Under the background of rapid economic development, the cultural industry has also made great progress, but the corresponding requirements are also put forward for the cultural industry. As the main competitiveness between countries, the development of cultural industry is of great significance to the development of cultural industry around the world. In the process of the development of cultural industry, it is first necessary to clarify the correlation between digital media and the development of cultural and creative industry, which plays a very important role in the development of cultural industry.

Keywords: Digital Media Art, Cultural and Creative Industry, Relevance.

1 INTRODUCTION

Under the background of increasing cultural resources, the development of digital media art has been promoted to some extent, and at the same time, the cultural and creative industry has also ushered in brand-new development opportunities [1]. For the cultural and creative industry, it consumes relatively little resources, and the degree of environmental pollution is relatively small. Due to its own advantage of high added value, it has a large space for development [2]. In addition, driven by digital media art, the development of the current cultural and creative industry has undergone corresponding changes. Based on this situation, this paper analyzes the relationship between digital media art and the development of cultural and creative industry.

2 AN OVERVIEW OF THE DIGITAL MEDIA ARTS AND CULTURAL AND CREATIVE INDUSTRIES

2.1 An Overview of Digital Media Art

Digital media art is a product produced by the combination of technology and art. It is mainly an emerging industry dominated by technology and supplemented by art. It is also an interdisciplinary discipline combining computer science and traditional movement. The main purpose of the development of digital media art is to cultivate talents with good scientific quality and artistic accomplishment, so that they can create innovative works of art through
computer software and corresponding design tools. Its characteristics are as follows: (1) media integration. Digital media art works are different from traditional works of art, and they are mainly created through computer technology and digital technology. For example, on the electronic map, the work fully embodies the integration of digital media art, but also adds color to people's lives. (2) Editable and replicability. For the art works created by computer multimedia, its production process needs to use some data information, and these data information can be edited and copied accordingly. But the works created in this way are often doubted by people, believing that this way buries the value of the art works to some extent. However, after the emergence of virtual reality, people have redefined the reality and virtual reality, making it the perfect combination of the editing and reproduction of real art and virtual art works. (3) Interaction and participation. With the development of the Internet, the distance between digital media art and the public has been shortened. People can appreciate and evaluate the works of art through the Internet platform, which has changed the traditional form of art appreciation, to a certain extent [3].

2.2 Overview of the Cultural and Creative Industries

Cultural creativity is a kind of industry produced by the combination of science and technology and art. It can integrate its own opinions and ideas into the artistic works. It can not only promote social development, but also create a kind of urban culture with unique artistic color. The characteristics are shown in Table 1.

<table>
<thead>
<tr>
<th>number</th>
<th>characteristic</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Uncertainty of the demand</td>
</tr>
<tr>
<td>2</td>
<td>Providers of creative products are very concerned about their own products</td>
</tr>
<tr>
<td>3</td>
<td>Need to integrate a variety of skills, creative product design elements are very rich</td>
</tr>
<tr>
<td>4</td>
<td>It is different and unique</td>
</tr>
<tr>
<td>5</td>
<td>Focus on vertical differentiation</td>
</tr>
<tr>
<td>6</td>
<td>Affected by the time factors</td>
</tr>
<tr>
<td>7</td>
<td>Continuity and long-term profitability</td>
</tr>
</tbody>
</table>

3 THE DEVELOPMENT STATUS OF DIGITAL MEDIA ART AND CULTURAL AND CREATIVE INDUSTRIES

3.1 The Development Status of Digital Media Art

As far as China is concerned, digital media art started relatively late, which makes a big gap between China's digital media art and other countries. However, due to the relatively rich cultural and artistic elements in China, coupled with the rapid development of China's hardware and software industry, it has laid a foundation for the development of China's digital media art. In addition, with the support of the Chinese government, China's digital media art has also ushered in a brand-new development opportunity. Especially during the "11th Five-Year Plan" period, it presented a leapfrog development mode, which not only promoted
the upgrading of the cultural industry structure, but also promoted the development of China's social economy to a certain extent [4]. However, compared with other countries, the market share of digital media art in China is relatively low, and it is relatively backward in terms of creativity, content and technology.

3.2 The Development Status of Cultural and Creative Industry In China

In recent years, China's creative industries and industry have developed rapidly with the support of the state, especially in Hong Kong and Taiwan. Of course, the development of Shanghai, Beijing and other regions has also achieved corresponding results, especially in the Shanghai region, they have taken their own advantages to seize the highland of the development of cultural and creative industry.

And also gradually established a variety of creative industry bases and parks.

4 THE CONNECTION BETWEEN DIGITAL MEDIA ART AND THE DEVELOPMENT OF CULTURAL AND CREATIVE INDUSTRIES

4.1 The Development of Digital Media Art Plays a Huge Role in Promoting the Cultural and Creative Industry

In the process of digital media art design, it contains relatively many design content, such as advertising design, visual design, etc., and more modern technologies are applied in the design process, such as computer technology, big data technology, etc. In the development process of cultural and creative industry, digital media art is usually taken as a booster. The main reason is that digital media art not only has a strong creative ability, but also has the corresponding comprehensive value, which promotes the leap-forward development of cultural and creative industry to a certain extent [5].

For the current cultural and creative industry, it involves a relatively wide range of fields, and is rich in content. Under the role of digital media art, it provides a corresponding carrier for cultural and creative products, making them more unique, and increasing the added value to a certain extent. For example, in the rise of recent years, the essence of the Taobao enterprise in recent years is to combine some cultural elements of the Palace Museum with modern cosmetics to produce cultural and creative products with Chinese cultural characteristics, as shown in Figure 1.

Figure 1 Palace Museum lipstick
In the process of designing such cultural and creative products, it is necessary to carry out the corresponding color design. The quality of the color is directly related to the quality of the leveling, especially the color in the ancient buildings has the corresponding cultural atmosphere, which can convey the symbolism of cultural information.

Generally speaking, the color of the object surface observed by the human eye is determined by the intensity of the three primary colors of red, rate and blue of the light in the incoming human eye. Generally, colors can be represented by the following formula:

\[ C = rR + gG + bB(0 \leq r \leq 1; 0 \leq g \leq 1; 0 \leq b \leq 1) \]

In formula: \( r, g, b \) represent the intensity of the three wavelengths of light, respectively, which is also called the three-stimulus value of color, mainly representing the RGB system of color. In the 24-bit computer display of the full color surface, with three bytes respectively, each value has 256 levels. All black \((0,0,0)\); white \((255,255,255)\), pure red \((255,0,0)\), green \((0,255,0)\), and pure blue \((0,0,255)\).

### 4.2 Cultural and Creative Industry Provides a Platform for Digital Media Art Design

As far as China's digital media art is concerned, its development is relatively late, and it has gradually entered into people's vision from 2000 \(^6\). At present, although its development process is only more than 20 years, but its development speed is relatively rapid, and now it has formed a variety of digital media art and design forms. For example, computer technology and multimedia technology combined form the Internet platform. Although these digital media arts are relatively backward compared with other countries, the cultural and creative industry of the Chinese market can develop rapidly due to the advantages of China \(^7\). Because of this, the demand for digital media art and design talents has also been increased. In order to meet the development needs of digital media art, most universities in China have gradually established the major of "digital media art design", and included it in the major of education technology. In addition, under the background of the expanding scale of the cultural and creative market, the public demand for cultural and creative products is also gradually increasing, and a large number of professionals are needed to support them \(^8\). Therefore, under the development of cultural and creative industry, it not only promotes the development of digital media art, but also improves the educational level of digital media technology to a certain extent. In general, the relationship between digital media art and design and the development of cultural and creative industry presents a complementary and common development relationship.

### 5 SPECIFIC APPLICATION METHOD OF DIGITAL MEDIA ART DESIGN IN THE CULTURAL AND CREATIVE INDUSTRY

#### 5.1 Adopt Digital Media Art and Design Means to Improve the Beauty of Urban Cultural and Creative Works

For cultural products, need to be through a variety of media channels to present it in the public view, this not only requires cultural products have the corresponding beauty, but also have a
certain three-dimensional sense, so as to deepen the public impression of cultural products, at this time, need to make full use of digital media art design means. For example, a factory in Australia, which is a commercial complex, has an aesthetic architectural design style, as shown in Figure 2. In the design process of the factory, the designer adopted the digital media art design means, to make the whole building bright in color, and also to depict a lot of artistic modeling, completely presenting the cultural characteristics of the Australian city.

![Figure 2 A special shopping mall in Melbourne, Australia](image)

5.2 Adopt Digital Media Art and Design Means to Improve the Economic Value of Cultural and Creative Works

In the process of cultural creative industry development, cannot leave the technical support of digital media resources, that is to say, need to use digital media art to create products, so as to enhance the artistic value of products, but also can optimize the structure of cultural creative industry, to promote the economic benefits of cultural creative industry has a very important role. For the current stage of China's film and television culture industry, its development has entered a stable state, which is all attributed to the digital media art [9]. Under the role of digital media art, the original concept of design can be presented through dynamic pictures, which not only effectively improves the aesthetic feeling of the picture, but also enhances the special effects of the picture to a certain extent [10]. In this way, the public can get a more complete visual enjoyment, which not only promotes the economic development of the film and television industry, but also promotes the development of the cultural industry to a certain extent. Therefore, in the development process of cultural and creative industry, great importance should be attached to the application of digital media art, so as to provide impetus for the development of cultural and creative industry.

5.3 Adopt Digital Media Art and Design Means to Improve the Innovation of Cultural and Creative Products

In the process of cultural and creative industry development, we need to pay attention to its innovation. Therefore, in the process of creating cultural and creative products, it is necessary to make full use of media art means, integrate some innovative methods and ideas into cultural and creative products, and present more innovative, creative and cultural content to the public, so as to enhance the innovation of cultural and creative products.

In the design process of some cultural and creative products, it is usually necessary to use digital media art design means and use different modules to present the design content. At this
time, each module is needed to be designed separately to meet the design needs of cultural and creative products for different modules. For example, in the design process of some environmental protection public service advertising, can be used by digital media art design into the shape of ice cream (as shown in figure 3), through the shape of ice cream to present the impact of climate environment changes on the earth, so not only can deepen the impression of the public, but also can inspire people's awareness to protect the environment.

Figure 3. Design of environmental protection public service advertisement

5.4 Adopt Digital Media Art and Design Means to Improve the Cultural Nature of Cultural and Creative Products

In the development process of cultural and creative industry, it is necessary to combine it with digital media art, so that cultural and creative products can have the corresponding interest and diversity, which plays an important role in improving the cultural nature of cultural and creative products.

Some dynamic works with cultural creativity can be designed through digital media art means. Compared with traditional static works, such dynamic works with cultural creativity can better reflect the cultural nature of the cultural industry.

For example, in the Mogao Grottoes in Dunhuang, the film "Dream Buddha Palace" was launched to the public through the digital exhibition center. In the creation process of the film, did we make full use of the digital media art and the current advanced technology to fully combine culture and art.

6 CONCLUSION

For now, the cultural creative industry development has made corresponding achievements, and social and economic structure has changed, in order to be able to further promote the development of our cultural creative industry, need to make full use of digital media art, and clear the correlation between the two, this to promote the development of our country cultural creative industry plays an important role.
REFERENCES

Research on the Dependency Distance of Quantifiers in Modern Chinese Based on Big Data Text Taking the Dot Quantifier “粒(particle)” as an Example

Linlin Zou
linlangyuxia@foxmail.com
College of Chinese Language and Culture, Jinan University, Guangzhou, China

Abstract: In the era of information explosion, quantifiers were often ignored in the past natural language processing (NLP). However, with the deepening of NLP system, the importance of quantifiers in the field of NLP, information retrieval, Chinese teaching, and corpus construction has become increasingly prominent. By analyzing its Nominal structure semantic categories, we can find that: category b (specific affairs) has the highest proportion, accounting for 74.98%. b4 (materials) account for 26.38% of b category; category f (social activities) is the least, appearing only twice, and its proportion can be ignored. Therefore, material subclass of specific affairs class is often used in combination with “粒”. In Liu’s paper, he believes that the mean dependency distance of Chinese is 3.662, so the average number of words that Chinese users need to remember when processing sentences is about 3. By calculating the entries in the corpus, we find that the distance between “粒(particle)” and the subject or topic of the sentence is about 3.217.

Keywords: Chinese, Dependency Distance, Quantifiers, 粒(particle), Big Data.

1 INTRODUCTION

The widespread use of quantifiers is an important feature that distinguishes modern Chinese from other languages of the Sino-Tibetan language family and Indo-European language family. In the research process of Chinese grammar history, Chinese quantifiers have always been the research object of linguists. In the 1950s, a quantifier was named by “暂拟汉语教学语法系统简述” (A Brief Introduction to the Tentative Chinese Teaching Grammar System): “The word that represents the quantitative unit of things or actions.”

With the development of linguistics and the deepening of research, the research on new perspectives and new theories of Chinese quantifiers is increasing. Chinese quantifiers, typology, cognitive linguistics and other fields of cross research are also emerging.
According to the research of linguist Dryer, modern Chinese (broadly speaking, modern Chinese refers to the language used by the Han nationality after the May 4th Movement in 1919) is regarded as an atypical VO language \(^1\). The focus of typological research is mostly on word order and the correlation between word order and other features. Quantifiers are not the focus of typological research, so that the importance of quantifiers in word order is often ignored. Until recently, Comrie \(^2\) and Haspelmath et al. \(^3\) began to take quantifiers as an important parameter in typology. At the same time, with the continuous development of natural language processing (NLP) systems, it is found that quantifiers play a crucial role in word order types in the language system where quantifiers are forced to be used. The division and classification of Chinese quantifiers have an important impact on the field of NLP, information retrieval, Chinese teaching, corpus construction, etc. Clarifying the division and classification of Chinese quantifiers and measuring the dependency distance between modern Chinese quantifiers and subjects will play an important role in the field of NLP, especially in ambiguity processing and parallel translation.

### 2 LINGUISTIC BACKGROUND

#### 2.1 Modern Chinese

The concept of modern Chinese has broad and narrow meanings. In a broad sense, modern Chinese refers to the language used by the Han nationality after the May 4th Movement. It not only includes modern standard Chinese (Putonghua), but also includes Chinese dialects. In a narrow sense, modern Chinese only refers to the common language of modern Han nationality--modern standard Chinese Putonghua.

#### 2.2 Quantifier

There are mainly two explanations about the causes of Chinese quantifiers: the first one comes from the functional school; The second interpretation comes from the typological school.

In many studies, people use “classifier” to translate Chinese “quantifiers”. However, some scholars believe that the category of classifiers is not equal to the meaning category of quantifiers. This is largely because the classification function of Chinese classifiers is not obvious, and the systematicness and motivation are not strong. Zhu pointed out in his analysis that sometimes there is a certain relationship in the meaning between nouns and the individual quantifiers that match them, such as something with extensibility is used for “张 (Zhang)” modification, which is a sort of classification \(^4\). However, it should be noted that Zhu added a sentence “This is only a few cases”, which shows that the motivation of Chinese classifiers is not very strong. For example, when weighing “鱼 (fish)”, “毛巾 (towel)” and “消息 (news)”, we use “条 (tiao)”, while “一只狗 (a dog)” and “一只狼 (a wolf)” use different classifications.

This paper believes that the connotation category of “classifiers” cannot cover the connotation category of Chinese quantifiers, so this paper still uses the old name of “quantifiers”, and divides Chinese noun quantifiers into four categories: individual quantifiers, collective quantifiers, capacity quantifiers and body metaphor quantifiers according to the theory of cognitive linguistics. The most numerous, frequently used and representative of Chinese noun quantifiers are the individual quantifiers. The study of Chinese individual quantifiers is also the focus of
linguists. More research on quantifiers based on big data will help us deepen our research on quantifiers and contribute to the development of Chinese NLP technology. This paper based on big data corpus mainly studies the dependency distance between Chinese quantifiers and subjects by taking quantifier “粒 (particle)” as an example. Examples are as follows:

- Nǐ shì yī lì jīnzi, zài nǎlǐ dōu huì fāguāng.
  You are a piece of gold that shines everywhere.
- Gèrén de tóngqíng xīn rútóng yī lì shāzi.
  Personal compassion is like a grain of sand.

2.3 Dependency Distance

Dependency distance is an important concept in the field of dependency grammar research. It refers to the linear distance (inserted reference) between two words with syntactic relations in a sentence. The size of dependency distance reflects the constraint of human cognitive mechanism on syntactic structure (inserted reference). Dependency grammar holds that dependency distance is the linear distance between the dominant word and the subordinate word in a sentence.

With the help of CoreNLP developed by Stanford University, the version used in this paper is 4.4.0. The dependency analysis tree with “粒” is as follows:

- Tā yī kǒu dōngxī dōu méi chī, chī jǐ lì huāshēngmǐ dé la!
  He didn’t eat a mouthful of food, and he just ate a few peanuts!

The following figure 1 shows the dependency structure tree of the example sentence.

![dependency structure tree of the example sentence.](image)

In this sentence, the noun “东西 (food)” is the nominal subject of the sentence. The numeral “几 (a few)” and the quantifier “粒” forms a quantitative phrase to modify the noun “花生米 (peanut)”.

We want to use the analysis of dependency grammar to calculate the distance between quantifier “粒” and the subject or topic in the sentence, so as to show the closeness of the relationship between them.
3 A STUDY ON THE QUANTITATIVE PROPERTIES OF QUANTIFIERS

This paper mainly relies on the BCC corpus (http://bcc.blcu.edu.cn/) of Beijing Language and Culture University to conduct a diachronic retrieval of the quantifier “粒 (particle)” in modern Chinese, and uses the dependency distance formula to calculate the dependency distance between the quantifier “粒 (particle)” and the subject. Diachronic retrieval of it in BCC corpus is shown in the following figure 2:

![Figure 2. Frequency Statistics of Quantifiers by years.](image)

In this section, we would discuss nominal structure (NS) behind it and quantitatively study its dependency distance with the subject or topic of the sentence.

3.1 A Brief Study on the Semantic Categories of “粒”

We counted the use of the NSs behind “粒” in the corpus, and summarized the NSs with frequency ranking within 1000. The following Table I shows the NSs of the first five and the last five and their frequency.

<table>
<thead>
<tr>
<th>NS</th>
<th>FO</th>
<th>F</th>
<th>NS</th>
<th>FO</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>沙 (sand)</td>
<td>1</td>
<td>296</td>
<td>阿斯匹林 (Aspirin)</td>
<td>996</td>
<td>1</td>
</tr>
<tr>
<td>米 (rice)</td>
<td>2</td>
<td>278</td>
<td>VC (Vitamin C)</td>
<td>997</td>
<td>1</td>
</tr>
<tr>
<td>种子 (seed)</td>
<td>3</td>
<td>203</td>
<td>排骨 (spareribs)</td>
<td>998</td>
<td>1</td>
</tr>
<tr>
<td>沙子 (sand)</td>
<td>4</td>
<td>170</td>
<td>牛肉干 (dried beef)</td>
<td>999</td>
<td>1</td>
</tr>
<tr>
<td>饭 (meal)</td>
<td>5</td>
<td>118</td>
<td>蝌蚪 (tadpole)</td>
<td>1000</td>
<td>1</td>
</tr>
</tbody>
</table>

Note: Frequency order (FO), Frequency (F)
We plotted the frequency and order of the NS behind “粒”， the following figure 3 is obtained.

![Figure 3. Statistics of NS frequency.](image)

In Figure 3, the abscissa is the order and the ordinate is the frequency. The number of the first 100 NSs are 3268, accounting for 59% of the total. That is to say, the words behind “粒” have a lot of aggregation characteristics, often those words, such as “沙 (sand), 米 (rice), 饭 (rice), 石头 (stone), 珍珠 (pearl)”, etc.

Obviously, the frequency of NSs behind “粒” shows a power law distribution. The fitting function is as follows:

\[ y = 619.91x^{-0.923} \quad (0 \leq x \leq 1000, x \in N^*) \quad (1) \]

Figure 3 and Formula 2 also show, to some extent, that it has a preference for NS, that is, it needs to match words of semantic categories within a certain category. Only those small words such as “沙子 (sand)” often appear, and can’t be collocated randomly, such as “一粒太阳 (a sun)”.

In order to explore the semantic category characteristics of NS, we need to use some semantic category tables in common domains. In the paper, “现代汉语分类词典” (A Thesaurus of Modern Chinese, TMC) is used as a reference semantic category table.

TMC is a dictionary that is classified and arranged according to the meaning of words. 82955 entries were collected in the TMC. It is developed based on the material warehouse and is also arranged according to the classification system of five semantic levels. There are 9 first level category, 62 second level category and 508 third level category. And there are 2057 in the fourth level category and 12659 in the fifth level category. The categories reflect the broad overview of the whole social life and Chinese vocabulary. In the meantime, TMC reflects the synonymous and antonymous relationship of words in detail. Therefore, TMC is a scientific semantic classification of modern Chinese \(^{[5-6]}\). So it was used to calculate the semantic category of quantifiers.

With the help of TMC, quantitative statistics of NSs semantic categories are shown in the figure 4 below:
In Figure 4, the letters a-g represent the largest semantic category names, respectively: biology, specific affairs, abstract things, space-time, biological activities, social activities, and auxiliary words. The number followed by each letter represents the sub category under this category.

Through the analysis of Figure 4, it is clear that category b has the highest proportion, accounting for 74.98%. b4 (materials) account for 26.38% of b category; category f is the least, appearing only twice, and its proportion can be ignored.

3.2 A Quantitative Study on the Dependency Distance of “粒”

In order to further explore the deep cognitive characteristics of “粒”, we use dependency grammar to calculate the distance between quantifiers and sentence subjects or topics.

This paper refers to Liu’s formula for calculating the mean dependency distance (MDD) of Chinese sentences [7]. The formula is as follows:

\[
MDD = \frac{1}{n-1} \sum_{i=1}^{n-1} |DD_i|
\]  

Because quantifiers are often combined with nouns to form NSs, most quantifiers cannot be directly controlled by the subject or theme of the sentence. Therefore, if we follow Liu’s formula, we cannot directly calculate the dependency distance between quantifiers and sentence subjects. But the NS of nouns is often dominated by the sentence subject, so we calculate the dependency distance between NS and the subject to represent the dependency distance between quantifiers and subjects.

In Liu’s paper, he thinks that the MDD of Chinese is 3.662, so the average number of words that Chinese users need to remember when processing sentences is about three. By calculating the entries in the corpus, we find that the distance between “粒” and the subject or topic of the sentence is about 3.217.
4 CONCLUSIONS

In the era of information explosion, we can access all kinds of texts every day, and we can also collect more and more texts. Then, after having such text data, it is necessary to do in-depth data processing on them. Therefore, it is necessary to conduct an in-depth study of the syntactic and semantic structure of the text. This paper selects the so-called stop words “quantifiers” that are often ignored in Chinese NLP research, and takes “粒” as an example to discuss the necessity of mining potential information of text based on big data.

Quantifiers were often ignored in the past natural language processing. However, with the deepening of NLP system, the importance of quantifiers in the field of NLP, information retrieval, Chinese teaching, and corpus construction has become increasingly prominent.

By analyzing its NS semantic categories, we can find that: specific affairs have the highest proportion, accounting for 74.98%. Materials account for 26.38% of specific affairs; social activities are the least, appearing only twice, and its proportion can be ignored. Therefore material subclass of specific affairs class are often used in combination with “粒”。In Liu’s paper, he believes that the MDD of Chinese is 3.662, so the average number of words that Chinese users need to remember when processing sentences is about 3 [7]. By calculating the entries in the corpus, we find that the distance between “粒” and the subject or topic of the sentence is about 3.217.

REFERENCES

Research on the Impact of Different Power Structures on the Supply Chain of Fresh Produce

Dandan Wang1*, Wei Gao1

*2465292480@qq.com

Business Travel Academy, Sichuan Agricultural University, Guankou Street, Chengdu, China1

Abstract: Considering the influence of fresh-keeping effort level and value-added service level on consumer demand in the sales process of fresh e-commerce platform, the consumer demand function under the joint influence of fresh-keeping effort level and value-added service level was constructed, and a single fresh food manufacturer and a single fresh e-commerce platform were studied (hereinafter replaced by retailers) constitutes the influencing factors of the secondary fresh e-commerce supply chain. By constructing the manufacturer-led and retailer-led fresh e-commerce supply chain decision model, the influence of freshness demand elasticity coefficient, value-added service elasticity factor and other factors on optimal decision-making is analyzed. The optimal decision of the two modes is compared. The results show that: 1) under the dominance of retailers and the elasticity coefficient of value-added services is low, the direct sales price and demand of the channel are the highest, but the preservation efforts provided by the manufacturer are higher, and the profit of the retailer will be much higher than the profit of the manufacturer; (2) When the manufacturer is led and the elasticity coefficient of value-added services is low, the overall profit of the supply chain is the lowest at this time.

Keywords: Different Power Structures, Freshness Preservation Efforts, Value-Added Services.

1 INTRODUCTION

With the rapid development of the Internet and the rapid popularity of e-commerce. As of 2020, China's fresh e-commerce market has developed rapidly, with the scale of the fresh e-commerce industry reaching 458.5 billion yuan, an increase of 64.0% over 2019. It is expected that fresh e-commerce will maintain rapid growth for a period of time in the future, and the scale of the fresh e-commerce industry will exceed one trillion yuan by 2023. In this context, it is of great significance to explore the cooperative relationship and optimal decision-making of value-added services and preservation efforts on the main parties of the supply chain, which is of great significance to improve the performance of the supply chain of fresh agricultural products.

The literature related to the research in this paper mainly focuses on two directions. The first research direction is the study of supply chain decision-making under different power structures. The questions they study vary from issue to issue. Wang et al. (2017) and Sun et al. (2020) consider the situation where different supply chain members have dual-channel supply chains. Choi (1991) and Shi et al. (2013) explore the impact of different power structures on supply chain members under different demand function forms. Zhang Guoxing et al. (2015), when
considering wholesale price as an exogenous variable, concluded that a single member of the supply chain can always benefit from market dominance, but for the entire supply chain channel, whether dominance is beneficial depends on the market potential comparison between the two channels. Liu et al. (2022) construct a two-tier transnational green supply chain model led by manufacturers and retailers, respectively, when considering tariffs, supply chain power structure, and consumers' green preferences. Li Xinran et al. (2018) constructed a closed-loop supply chain system composed of individual manufacturers and retailers under the consideration of free riders, and explored the influence mechanism of different power structures on the decision-making of closed-loop supply chain system.

The second research direction is the study of the dual effort factors of the supply chain. Bai Shizhen et al. (2018), Liu Molin et al. (2020) considered the impact of fresh-keeping efforts and promotion efforts, and fresh-keeping efforts and value-added services on fresh e-commerce supply chain decisions. Zhang Chong, Liu Ying et al. (2021) consider the quality efforts of manufacturers and the sales efforts of retailers in the model. Hua (2001) takes into account the advertising promotion method in the model.

The main differences between this paper and the aforementioned literature are: (1) the impact of fresh-keeping effort coefficient and value-added service coefficient on supply chain decision-making under different power structures is studied, while existing studies have not considered the dual impact of fresh-keeping effort and value-added service together; Or do not consider different supply chain power structures; (2) With the popularity of the Internet and the rapid development of self-media, product pricing information has become more and more transparent, so this article takes wholesale price as an exogenous variable.

### 2 DESCRIPTION OF THE PROBLEM AND PARAMETERIZATION

This paper considers a supply chain consisting of a fresh manufacturer (M) and a retailer (O), in which the fresh manufacturer is responsible for activities such as the production, packaging and other activities of the fresh product, that is, the cost per unit of fresh product produced is $c$; The retailer is responsible for selling fresh products at a price and providing consumers with relevant value-added services $e_2$, including promotions and content services related to fresh products [11], and after the consumer places a successful order, the retailer is responsible for delivery, and the supplier provides freshness preservation measures.

This paper expresses the demand function of fresh products as:

$$d = 1 - p + \alpha \theta_0 e_1 + \beta e_2$$

where $d$ represents the demand of the channel, $\alpha$ ($\alpha > 0$) represents the freshness elasticity coefficient, $\theta_0$ represents the initial freshness of fresh product distribution, $e_1$ ($e_1 > 0$) represents the level of freshness preservation effort, $\beta$ ($\beta > 0$) represents the elastic coefficient of value-added services, and $e_2$ ($e_2 > 0$) represents the value-added service level $\alpha > \beta$.

Suppose the manufacturer's cost of freshness and the retailer's cost of providing value-added services are respectively:
where \( C(e_1) \) represents the manufacturer's freshness cost, \( k_1 > 0 \) represents the manufacturer's product freshness cost coefficient; \( C(e_2) \) represents the cost of value-added services paid by retailers, and \( k_2 (k_2 > 0) \) represents the retailer's value-added service cost factor.

In summary, the profits of manufacturers and retailers in the secondary supply chain are respectively

\[
\pi_m(e_1) = (w - c)d - \frac{1}{2} k_1 e_1^2 \\
\pi_o(p, e_2) = (p - w)d - \frac{1}{2} k_2 e_2^2
\]

where \( \Pi_m \) represents the manufacturer's profit, \( \Pi_o \) represents the retailer's profit, and \( w \) is an exogenous variable that represents the wholesale price determined by a long-term contract.

For clarity, the symbols used in this model are specified as follows: the superscript * in the symbol represents the optimal solution (or optimal target value) of the corresponding optimization model. The meaning of symbols other than these will be explained in the text.

3 SUPPLY CHAIN DECISIONS UNDER DIFFERENT POWER STRUCTURES

3.1 Manufacturer-led supply chain decisions (M)

In the M model, the decision order is: the manufacturer first determines the freshness of the fresh produce \( e_1 \), then the retailer decides the retail price of the product \( p \) and the value-added service \( e_2 \), and the profit of the manufacturer and retailer is (4) (5) respectively.

Theorem 1 Manufacturer-led, when \( 2k_2 - \beta^2 > 0 \), by reverse induction, the optimal equilibrium decision between supplier and retailer is:

Optimal freshness preservation efforts:

\[
e_{1}^{*} = \frac{(w - c) \alpha \theta_0}{k_1 (2k_2 - \beta^2)};
\]

Optimal value-added services for retailers:

\[
e_{2}^{*} = \frac{\beta + \alpha \beta \theta_0 e_1 - \beta w}{2k_2 - \beta^2};
\]
Optimal retail price for retailers:

\[ p^* = \frac{k_1k_2\left(2k_2 - \beta^2\right) - ck_2^2\alpha^2\theta_0^2 + w\left[k_2^2\alpha^2\theta_0^2 + (k_2 - \beta^2)\left(2k_2 - \beta^2\right)k_1\right]}{k_1\left(2k_2 - \beta^2\right)^2}; \]

The total demand is:

\[ d^* = \frac{k_1k_2\left(2k_2 - \beta^2\right) - ck_2^2\alpha^2\theta_0^2 + w\left[k_2^2\alpha^2\theta_0^2 - k_2k_2\left(2k_2 - \beta^2\right)\right]}{k_1\left(2k_2 - \beta^2\right)^2}; \]

The total profits of manufacturers and retailers are:

\[ \pi_m^* = \frac{k_1k_2\left(2k_2 - \beta^2\right)(w-c) + c^2k_2^2\alpha^2\theta_0^2 - 2w\left(k_2^2\alpha^2\theta_0^2 - k_2k_2\left(2k_2 - \beta^2\right)\right) + w\left(k_2^2\alpha^2\theta_0^2 - 2k_2k_2\left(2k_2 - \beta^2\right)\right)}{2k_1\left(2k_2 - \beta^2\right)^2}; \]

\[ \pi_o^* = \frac{k_1\left(2k_2 - \beta^2\right)^2 - ck_2\alpha^2\theta_0^2 + w\left(k_2\alpha^2\theta_0^2 - k_1\left(2k_2 - \beta^2\right)\right)^2}{2k_1^2\left(2k_2 - \beta^2\right)^3}. \]

3.2 Retailer-led supply chain decisions (O)

In the O model, the decision order is: the retailer first determines the retail price of the product \( p \) and the value-added service \( e_2 \), the supplier determines the freshness of the fresh produce \( e_1 \), and the profit of the manufacturer and retailer are equation (4) (5) respectively.

Theorem 2 Under manufacturer dominance, when \( 2k_2 - \beta^2 > 0 \), the optimal equilibrium decision between manufacturer and retailer is:

Optimal freshness preservation efforts of manufacturers: \( e_0^* = \frac{(w-c)\alpha\theta_0}{k_1}; \)

Optimal value-added services for retailers:

\[ e_2^* = \frac{\beta\left[k_1 + w(\alpha^2\theta_0^2 - k_1) - c\alpha^2\theta_0^2\right]}{k_1\left(2k_2 - \beta^2\right)}; \]

Optimal retail price for retailers:

\[ p^* = \frac{k_2\left(k_1 - c\alpha^2\theta_0^2\right) + w\left(k_2\alpha^2\theta_0^2 + k_1\left(2k_2 - \beta^2\right)\right)}{k_1\left(2k_2 - \beta^2\right)}; \]
The total demand: \( d^{\alpha'} = \frac{k_1}{k_1} \left[ k_1 - c_0 \alpha^2 + w \left( \alpha^2 - k_1 \right) \right] \).

The total profits of manufacturers and retailers:

\[
\pi_m^{\alpha'} = \frac{2k_1k_2(w - c) + 2wc \left( k_1k_2 - \beta^2 \alpha^2 \right) - w^2 \left( 2k_1k_2 - \beta^2 \alpha^2 \right) + c_0^2 \beta^2}{2k_1 \left( 2k_2 - \beta^2 \right)}.
\]

\[
\pi_0^{\alpha'} = \frac{k_1 \left[ k_1 - c_0 \alpha^2 \right] + w \left( \alpha^2 - k_1 \right)^2}{2k_1 \left( 2k_2 - \beta^2 \right)}.
\]

4 COMPARATIVE ANALYSIS AND CASE ANALYSIS

In this section, the impact of different power structures in the supply chain on direct prices, freshness preservation efforts and value-added services, demand, and revenue of supply chain members is discussed.

Firstly, by comparing and analyzing the impact of the power structure of different supply chains on the direct selling price decisions of supply chain members, the following conclusions can be drawn.

Proposition 1 In the M game and the O game, the direct selling price under the M game is greater than the direct selling price under the O game. Namely \( p^M > p^O \).

According to proposition 1, it can be seen that the direct sales price in the O model is greater than the direct sales in the M model; The reasons and mechanisms can be explained as: retailers have greater initiative when they are the dominant players, so retailers will set direct sales prices that are conducive to the growth of their own interests; In the M model, manufacturers do not want retailers to set excessively high retail prices in order to increase consumer demand and ensure their own profits.

Secondly, by comparing and analyzing the impact of power structures in different supply chains on manufacturers' freshness preservation efforts and retailers' value-added service decisions, the following conclusions can be drawn.

Proposition 2

(1) when \( 0 < \beta < \sqrt{k_2} \), \( e_1^M < e_1^O \); when \( \sqrt{k_2} < \beta < \sqrt{2k_2} \), \( e_1^M > e_1^O \);

(2) \( \frac{\partial e_2^M}{\partial \alpha} > 0 \), \( \frac{\partial e_2^M}{\partial \beta} > 0 \); \( \frac{\partial e_2^O}{\partial \alpha} > 0 \), \( \frac{\partial e_2^O}{\partial \beta} > 0 \).

According to the (1) of proposition 2, the following explanation can be made: under the O model, retailers will reduce the cost of value-added services, and then hope that suppliers will
improve the freshness of fresh products, thereby stimulating consumers' demand for fresh products, and retailers will get more profits from them, but under the M model, improving the freshness of fresh products requires suppliers to invest in preservation costs, and their profits will decline, so suppliers will not invest too much preservation efforts to ensure their own profits. From (2), it can be seen that whether it is M or O model, the supplier's freshness preservation effort is positively correlated with the freshness preservation effort coefficient and the value-added service coefficient. This phenomenon can be explained as: in the M or O model, whether the supplier improves the freshness preservation effort or the retailer improves the value-added service, the supplier's freshness preservation effort can be improved. Thirdly, by comparing and analyzing the impact of different supply chain power structures on demand decision-making, the following conclusions can be drawn.

\[ \text{Proposition 3 when } 0 < \beta < \sqrt{k_2 - \frac{c}{w - c}}, \quad d^{M'} < d^{O'} ; \text{when } \sqrt{k_2 - \frac{c}{w - c}} < \beta < 1, \quad d^{M'} > d^{O'}. \]

According to proposition 3, when consumers are not particularly sensitive to the value-added services provided by retailers, the total demand for the supply chain under the O model is greater than that under the M model; When consumers are sensitive to the value-added services provided by retailers, the total demand for the supply chain under the M model is greater than under the O model, which is obviously consistent with the facts. This is because when consumers' sensitivity to value-added services is \( \left[ 0, \sqrt{k_2 - \frac{c}{w - c}} \right] \) within the range, retailers will provide higher value-added services to attract more consumers to buy fresh products, and the increase in demand brought by value-added services is greater than the decrease in demand caused by the increase in direct sales prices led by retailers, and the total demand of the final supply chain shows a growing trend. When consumers' sensitivity to value-added services is \( \left[ \sqrt{k_2 - \frac{c}{w - c}}, 1 \right] \) within the range, although consumers attach more importance to the value-added services provided by retailers, as can be seen from proposition 1, the direct sales price under the M model is the lowest, so the demand led by suppliers is the largest.

Finally, by comparing and analyzing the impact of power structures in different supply chains on manufacturers' profits and retailers' profits, the following conclusions can be drawn.

\[ \text{Proposition 4 (1) } \pi_m^{M'} > \pi_m^{O'} ; \text{(2) } \frac{\partial \pi_m^{M'}}{\partial \alpha} > 0, \quad \frac{\partial \pi_o^{O'}}{\partial \beta} < 0. \]

According to (1) of proposition 4, it can be seen that in the M model, the manufacturer is the leader and has absolute power in the supply chain, so the supplier will make a series of decisions to maximize its profits. It can be seen from (2) that in the retailer-led model, the profit of retailers increases with the increase of the manufacturer's freshness preservation coefficient, and
decreases with the increase of the retailer's value-added service coefficient. This phenomenon can be explained by the fact that in the case of supplier-led, although the cost of freshness preservation efforts continues to increase, suppliers will still increase the freshness of fresh agricultural products in order to obtain more profits, thereby expanding the demand for channels, and retailers' profits will increase; The greater the cost of value-added services borne by retailers, the lower their profits.

Due to the limitation of space, in this chapter, this paper only discusses the impact of freshness demand elasticity and service demand elasticity parameters on the profits of fresh suppliers, and this paper sets the relevant parameters based on the model assumptions, setting the parameters: \(c=0.1, \ w=0.5, \ a=5, \ b=4, \ \theta_0=0.9, \ k_1=20, \ k_2=25.\)

It can be seen from Figure 1 that the increase in \(a\) has a positive effect on the profit of suppliers under the M/O mode, but under the M model, the impact of \(a\) on the profit of suppliers is more significant. As can be seen from Figure 2, when the \(b\) is small, the profit of the supplier in the M mode is larger; When the \(b\) is large, the opposite is true, and these situations are similar to our conclusions above.

5 EPILOGUE

In this paper, the optimal decision-making under the two supply chain structures is compared and analyzed in a two-tier supply chain consisting of a fresh food manufacturer and a retailer. It
is found that: (1) under the dominance of retailers and the elasticity coefficient of value-added services is low, the direct sales price and demand of the channel are the highest, but the preservation efforts provided by the manufacturer are higher, and the profit of the retailer will be much higher than the profit of the manufacturer; (2) When the manufacturer is led and the elasticity coefficient of value-added services is low, the overall profit of the supply chain is the lowest at this time, so in this case, the manufacturer should delegate authority to the retailer.

It should be pointed out that there are many research directions that can be extended in this paper. For example, a simple secondary supply chain can be extended to multiple manufacturers, multiple retailers, and a cost-sharing contract can be added to determine a threshold that optimizes the profits of both manufacturers and retailers.

REFERENCES


A Research on the Development of a Publicly Owned Natural Assets Management and Supervision Platform

Jiaqiang Ren¹a*, Huijing Chu¹b

* Corresponding author: 1939251760@qq.com, Chj20211202@163.com

School of Management, Shenyang Jianzhu University, Shenyang, Liaoning, China

Abstract: To meet the practical needs of natural resource management and supervision, based on the inventory results of natural resource assets such as the state-owned land, minerals, oceans, forests, grasslands, wetlands, water, etc., and the basic information platform for the land and space, and by the unified provincial standards, we will gradually build an integrated natural resources management and supervision platform with functions including statistical accounting, asset allocation, and evaluation and assessment through exploring the measures such as "business reshaping, data integration, and application integration", etc. The platform will provide strong and favorable technical support for planning and preparation, administrative review and approval, monitoring and supervision, evaluation and decision-making, etc. for the natural resources, and effectively improve the level of modernization of natural resource management.

Keywords: Natural Asset Management, Regulation, Platform Development.

1 INTRODUCTION

There are many types and large quantities of natural resources. Therefore, their asset value assessment is affected by multiple factors such as resource type, environment, market value, and ecological benefits. Their management and value assessment is quite complex and calls for theories, methods, and norms in the work. Since the 18th National Congress of the CPC, the institutional mechanism for the management of natural resource assets owned by the public has been gradually established and improved, which has played a positive role in promoting the construction of ecological civilization. However, there are still problems such as a lack of owners in place and unclear management responsibilities. The problems of excessive resource exploitation, weak protection, and ecological degradation are still quite serious. Currently, the management of natural resource assets in various regions is still in the exploration stage. Standards and measures with uniform regulations have not yet been formed, and the informatization construction is insufficient. Establishing an information support platform to serve natural asset management is demanded if the owner's responsibilities are to be fulfilled, and it is also demanded if the owner's rights and interests are to be realized [1].
Following the general principle of integrated management of natural resources, we analyzed the characteristics of various work tasks and businesses, studied the key technologies such as efficient big data computing, platform cloud architecture, efficient spatial data rendering, and visualization components, and built a natural resource assets management information system and carried out application practice to support natural resource assets management and supervision.

2 SYSTEM REQUIREMENTS ANALYSIS

2.1 Challenges for the Natural Resources Information Management

Natural resource asset management is a systematic project that involves a wide range of areas and multiple sectors. It is comprehensive and systematic. Traditional management methods are often inadequate, and a unified management and supervision platform has not been established between different departments, making it difficult to effectively conduct resource management. As a result, it faces five outstanding problems including a lack of owner in place, unclear assets inventory, insufficient market allocation, imperfect revenue management, and inadequate ownership assessment and supervision.

2.2 The Necessity of System Building

The integrated natural resources management and supervision platform provides favorable technical support for planning and preparation, administrative examination and approval, monitoring and supervision, and evaluation and decision-making for natural resource development and utilization, and effectively improves the modernization level of natural resource management. In the background of the “Internet +” era, it is imperative to establish an integrated natural resource management and supervision platform to form unified data standards and norms and break down departmental barriers.

3 SYSTEM DESIGN

3.1 System Architecture

The system architecture consists of an infrastructure layer, a data resource layer, a platform support layer, and an application layer. It is shown in “Fig 1”.
3.1.1 Infrastructure Layer

As a supporting environment, the infrastructure layer provides the infrastructure supporting the databases and application systems, including required computing resources, storage resources, network environment facilities, and server virtual systems. The infrastructure layer relies on the results of natural asset inventory and information platforms. It is established based on the government extranet interconnected by provincial, municipal, and county departments, which ensures the stability of operation and the convenience of use at the city and county levels [1].

3.1.2 Data Resource Layer

The prerequisite for centralized management of natural resource data at the data resource level is the formation of a unified data standard. Then relying on data management and service platforms and various data processing tools, we may continuously update, maintain, and manage natural resource data. In the natural resources integration database, the province's land, minerals, oceans, forests, grasslands, wetlands, water, and other planning data from different years are collected to ensure that all relevant data are further improved during the implementation process, and that data support is provided for higher level applications.
3.1.3 Platform Support Layer

The platform support layer completes various supporting tasks such as data collection, extraction, integration, quality inspection, operation, and maintenance based on the database, and enables data extraction, fusion, analysis, and computation by customizing basic algorithms and building computational models. The service interface solves the problems of decentralized storage and separate management of a large number of natural resource business data and provides the application system with the required spatial computing capabilities and public data support.

3.1.4 Application Layer

The application layer uses the basic data related to the daily management of natural resource assets and the index data of rich information and various forms to provide users with a visual page. It facilitates the end users to clearly and intuitively understand the situation of natural resource assets in the whole province. It has the function of directly realizing the work and serving the management work of natural resource assets.

3.2 Business Reshaping

When building an integrated management and supervision platform for natural resources, the business process should be clarified. Based on the needs of natural resource management and supervision, scientific management concepts and advanced and feasible information technology should be used to fully prepare and optimize the process, and the natural resource management business should be packaged into different categories and re-integrated to make the processing more flexible and the operation more convenient and faster. First of all, the core tasks of natural resources management and supervision should be clarified, for example, what functions should the natural resources integrated management information system and supervision platform have and what aspects should be covered. The relationship between various businesses should be clarified to form a business system. We need to comprehensively consider the business content of the natural resources such as the state-owned land, minerals, oceans, forests, grasslands, wetlands, water, etc., determine the relationship context, build an integrated management and supervision platform, establish a unified login portal and information entry and output channels, etc., improve the business process, strictly standardize and supervise the handling and implementation of each work task, and form a standardized manual covering processing procedures, business rules, required materials, time of limit of work, etc. Only by doing so will we be able to provide guidance and regulation to the processing of the related business.

3.3 Data Fusion

For the intensive management and supervision of natural resources, resource sharing, and business collaboration, forming a unified database system is a prerequisite. Based on the inventory results of national land, minerals, oceans, forests, grasslands, wetlands, water, and other natural resource asset and territorial spatial basic information platform data, we gathered and integrated diverse and heterogeneous natural resource-related data, clarified the data context and centralized storage, established management standards, and achieved the purpose of hierarchical data management. In the data fusion process, the target layer is treated according
to the thematic data layer, such as spot breaking, value assignment, area adjustment, spot merging, etc., to perfectly meet the data extraction requirements of various resources. Finally, it provides functions such as data conversion in various formats, automatic real-time data updates, annual data changes, and multi-party sharing.

3.4 Application Integration

To establish an intensive platform for the management and supervision of natural resource assets owned by the public, it is necessary to make full use of the existing hardware and software infrastructure and configuration, such as network environment, application system, information platform, etc., to carry out application integration and change the traditional way of resource management. We need to unify the portal to enable the full coverage of the natural resource asset management information system business, enable the unified scheduling of information and data, break down the barriers between departments, and strengthen the communication and coordination between departments with the help of information technologies such as cloud computing, big data and the Internet of Things, to have smooth communication channels at the departmental level. Secondly, the Internet + supervision method is adopted for unified supervision and management to ensure that various decisions are made in a smooth and orderly manner.

3.5 Highlights Of The Technology

3.5.1 Architecture Upgrade

Unlike the previous chart, cloud architecture is used to support the scheduling of the elasticity of the clusters and nodes. Nodes can be added and reduced to achieve resource controllability, and multiple nodes can be parallelized to improve system performance and computing efficiency with stronger performance stability and scalability of the cloud architecture cluster management. This is shown in “Fig 2”.

3.5.2 Service Model Upgrade

The service management system adopts multi-level decentralized management, one-level deployment, multi-level use, and information exchange for the service upgrades, as shown in “Fig 3”.

Fig 2 Platform Cloud Architecture
3.5.3 Dynamic Fast Rendering of Data

The massive vector data is directly displayed in an unsliced manner, saving a lot of time for image resizing and enabling browsing in seconds under high concurrency, achieving a browsing speed comparable to that of the tiles. When related data changes, they are updated synchronously on the client side, so that the data changes and information can be instantly shared.

3.5.4 Efficient Big Data Computing

The platform supports multi-source data access from various data sources such as MapGIS, SDE, and MapGisDatastore. Distributed and efficient computation can realize distributed spatial computation of hundreds of millions of vector data. Compared with the traditional computation model, the computation has been greatly enhanced and the computation results are directly presented in the form of rich statistical charts, providing visual support for decision-making.

3.5.5 Full Sharing of Natural Resource Information

A series of service resources such as the cloud service portal includes the application service system, and the relevant departments can apply for online access through the portal. The basic data, examination and approval, planning, evaluation, supervision, and other applications of the natural resource assets such as state-owned land, forests, oceans, grasslands, minerals, wetlands, and water are all incorporated into the integrated platform of natural resources.
management, which enables the whole chain business, assists daily office work and improves work efficiency.

4 CONTENT OF THE CONSTRUCTION

4.1 Database Construction

Natural asset management is complex work, and the complete, detailed and effective natural asset data is the premise and basis for smooth management and supervision. The natural asset database is built based on the inventory results of natural resource assets such as state-owned land, minerals, oceans, forests, grasslands, wetlands, water, etc., and is a basic information platform for territorial space. The database construction is completed following the requirements of a unified spatial positioning benchmark, unified naming rules, and a unified data format of the data structure.

4.2 Construction of the Information Release Platform

The natural asset data system is relatively quite huge. Achieving rapid display in the system not only places high requirements on management functions but also places high requirements on system performance. This platform applies Web service technology and SOA system architecture completes platform construction concerning OGC standard service specifications, realizes distributed management, centralized sharing, and service of natural asset information, supports multi-source heterogeneous data, supports third-party services based on OGC standards, supports various file structures, supports dynamic tile acquisition and other technologies, and facilitates the quick browsing and access by the users.

4.3 Data Center Management System Construction

To ensure quality data storage, preprocessing such as inspection, editing, and data conversion must be performed on the data by the relevant specifications to ensure the quality and standardization of data storage. The data center management system consists of a data management maintenance subsystem, a data exchange and update subsystem, a data query and statistical analysis subsystem, an application service management subsystem, and a system maintenance subsystem.

4.4 Construction of Data Sharing and Exchange Service System

The data information sharing service system is mainly built to complete the dynamic release of data information, information query, data sharing, and other functions, publishing the data information on the extranet portal website, making full use of and integrating natural asset database resources, building a public information service platform integrating spatial data and attribute data, realizing relevant data sharing, information browsing, and inquiry functions, providing information services to enterprises, the public and government departments, and improving the online approval and declaration system.
5 FUNCTIONAL ANALYSIS

The integrated natural resources management and supervision platform is an intelligent management platform that can perform data analysis, plan difference analysis, assist natural government approval, and serve natural resource supervision and regulation, the natural resource information sharing with the integrated functions such as statistical accounting, asset allocation, evaluation, and assessment. The main functions are as follows:

5.1 Data Analysis

Based on the data resource catalog, we collected and sorted out all kinds of territorial spatial data at all levels, established a data resource classification system, formed a natural resource data center, and provided auxiliary planning preparation and review tools for planning preparation organizations and personnel. The data center enables editing and revision of the planning data, quality monitoring of planning results, spatial analysis and evaluation, comparison and coordination of planning differences, and the three-district and three-line inspection of spatial planning zones to improve the planning preparation efficiency and quality of results. Through the establishment of an evaluation model, a comprehensive analysis of the current situation, planning, and socio-economic data existing in the database was carried out, and the carrying capacity and development suitability of the land and space resources was monitored and evaluated. Thus, a scientific basis is provided for the rational regulation and control of ecology, agriculture, urban space, etc[4].

5.2 Planning Difference Analysis and Compliance Review

Planning difference analysis, compliance review, and comparative analysis between multiple layers can be used for conflict detection in land planning, urban and rural planning, forestry planning, etc., to provide technical support for resolving various planning conflicts, conduct compliance reviews of each resource application, achieve a “speaking-with-data” method to determine whether the application conforms to the master plan and issue monitoring and analysis reports and compliance details to assist planning approval.

5.3 Auxiliary Natural Resources Government Approval

Functions such as graphic positioning and query services provide visual spatial graphic support and information query services for natural resource e-government approval. In addition, by spatially superimposing, caching, and statistical processing of various kinds of natural resource data, the top ten administrative review functions are enabled, and effective natural resource management can be carried out[5].

5.4 Serving the Regulation and Supervision of Natural Resources

For land resource supervision, the platform enables the whole life cycle supervision of land "pre-examination, approval, land supply, land use, compensation, and investigation", and provides overall supervision of land use pre-examination, construction land, land supply projects, development and consolidation projects, etc. For mineral resource supervision, the platform enables the full life cycle supervision of mineral resources, provides exploration
project registration, mining right application registration query, and supports multi-condition combination query, keyword query, etc. According to the specific needs, it also serves the regulation and supervision of other natural resources such as forests, wetlands, oceans, water, etc.

5.5 Serving Natural Resources and Information Sharing

Through service cluster management, it supports dynamic resource scheduling and elastic scaling; supports the registration, publication, and formation of service resource pools of vector spatial data, structural data, file type data, and three-dimensional models. Through the unified service portal, it provides a service resource center for government departments, corporations, and the public, supports self-service applications and offline use, and enables multi-party sharing of natural resource information.

6 PRACTICAL APPLICATIONS

6.1 A New Data Resource Classification System

Natural resource data has numerous sources and abundant volume. Standardized data management is required in building the new data resource classification system, management mechanism, and organization method for basic information, data source information, data dictionary, and storage information. According to the data resource catalog, various types of natural resource data are collected and collated at all levels to form a natural asset big data center. As shown in “Fig 4”.

![Fig 4 Resource Classification System Interface](image-url)
6.2 Analysis and Evaluation

By establishing a comprehensive evaluation model, the current situation, planning, and socioeconomic data existing in the database are comprehensively analyzed for weighted calculations, and the state of natural assets is tested and evaluated.

6.3 Auxiliary Natural Resources E-Government Approval

By displaying various types of natural resource data, the platform provides visual spatial graphic support for natural resources e-government approval. As shown in “Fig 5”.

![Fig 5 Natural Resources E-Government Approval Interface](image)

6.4 Supervision of Natural Assets

Taking state-owned land as an example, it realizes full-lifecycle supervision of land “pre-examination, approval, land supply, land use, land supplementary, and inspection”, and provides overall supervision of land use pre-examination, construction land, land supply projects, development and consolidation projects.

6.5 Information Resource Sharing

Through service cluster management, it supports dynamic resource scheduling and elastic scaling; supports registration and publication of various data and 3D models to form a data resource pool. Government departments, enterprises, and the public can apply for online use of the platform through the unified service portal. As shown in “Fig 6”.

![Fig 6](image)
7 CONCLUSIONS

Based on the natural resource asset inventory results, the spatial data of the territorial spatial basic information platform, and the natural resource management requirements, this paper constructs an integrated natural resource information management and supervision platform with the support of emerging technological means such as big data, the Internet, GPS, etc., to serve natural resource management and supervision. It enables real-time tracking of natural resource dynamics, breaks down departmental barriers, improves work efficiency, and effectively improves the modernization level of natural resource management capabilities. First of all, natural resource assets are a huge system, and there are many types of data. There is still a long way to go before we can achieve complete storage of data. Next, we need to carry out further data mining work and complete the supplementary storage of data. Second, it is necessary to continuously improve and maintain the system during actual application.

REFERENCES

[5] Innovating a “hardcore” national production of land and space basic information platform to empower the digital transformation of natural resources - China University of Geosciences Asset Management Co., Ltd. (cug.edu.cn)
The Spatial Spillover of Digital Economy on Green Innovation: Evidence from China

Yuming Zhai, Ning Xie*
*206121158@mail.sit.edu.cn
School of Economics and Management, Shanghai Institute of Technology, Shanghai, China


Keywords: Digital Economy, Green Innovation, Spatial Effect.

1 INTRODUCTION

As China has entered the era of Industry 4.0, the digital economy under the rapid development has now become the backbone of promoting the high-quality development of the domestic economy. This is conducive to helping enterprises transform the mode of economic development and promote the construction of ecological civilization [3]. Meanwhile, green development has become a new global development trend. China responded quickly and positively. It put forward the concept of high-quality economic development including green and innovation, and formulated a nationwide action plan for energy conservation and emission reduction. As the combination of innovation driven strategy and green development strategy, green innovation has become one of the core contents of promoting sustainable economic development and the key path of building a community with a shared future for mankind. Therefore, how to give full play to the power of digital economy to promote green innovation has gradually become the focus of all sectors of society and government.

Green innovation refers to the technological innovation of green products and processes to reduce the environmental burden or achieve the goal of ecological sustainability. The digital economy is an economic activity with data as its resource, which promotes profound changes in the production and operation modes of enterprises and society. Based on resource-based theory, the digital economy provides new digital resources, platforms and development space for green innovation [5].
From the perspective of cost, the digital economy has the advantage of low cost, which improves the main body’s green innovation willingness. The digital economy breaks through the restriction of geographical distance through efficient information transmission [4], reduces the daily operating costs of enterprises, and increases the investment in green innovation. Secondly, the construction of digital infrastructure has improved the level of regional informatization, it can effectively promote the integration of related industries, optimize regional industrial layout. It also provides a management medium for government supervision, and promote the industrial structure to digital, rational and green transformation and upgrading (Kohli & Melville, 2018).

From the perspective of resources, digital technology can promote resource matching and green transformation of enterprises. Digital technology provides an information platform for enterprises to grasp market trends more quickly, respond to market demands in time, and improve enterprise resource matching. In addition, more rapid information transmission and richer access to knowledge will make the market environment more open and transparent. Enterprises must ensure their own survival and development through innovation, and accelerate the innovation of green products and processes [6].

Because the digital economy can achieve efficient information transmission in different regions, promote close cooperation between supply chain enterprises, and promote green innovation activities among regions. Therefore, we propose the hypothesis as follows:

Digital economy can improve urban green innovation and drive the coordinated development of adjacent cities.

2 Materials and Methods

2.1 Model Settings

The following model is built for the spatial spillover effect of digital economy on green innovation (Formula 1), which is the spatial Doberman model (SDM).

\[
\text{Lngi}_{t} = \alpha_{0} + \rho \text{Wln} \text{dei}_{t} + \phi_{1} \text{WLnde}_{t} + \alpha_{1} \text{ln} \text{dei}_{t} + \phi_{2} \text{C}_{t} + \mu_{t} + \delta_{t} + e_{i,t} \\
\]  

(1)

where \( W \) is the spatial weight matrix, \( \rho \) is the spatial autoregressive coefficient, and \( \phi_{1} \) and \( \phi_{2} \) is the coefficient of spatial interaction term.

2.2 Variables Description

2.2.1 Explanatory Variables

Digital economy (DE). Referring to the research of Zhao et al. (2020)[7], we selected five indicators: number of employees who engaged in software, mobile phone users number, the Internet broadband access users number, the telecom business income and the urban digital inclusive financial index. The five indexes are integrated into a comprehensive index through principal component analysis, which is recorded as De.
2.2.2 Explained Variable

Green innovation (Gi). Referring to the research of Liang et al. (2022)\(^2\), we choose the green invention patent application number as the index to measure green innovation. Due to the uneven development level of green innovation among cities, the number of green creation applications in some regions is 0, which may affect the subsequent calculation. Thus, we use the logarithm of the urban green patents number plus 1 as the dependent variable, which is recorded as Gi.

3.2.3 Control Variables

Referring to relevant literature, we selected five control variables, which can reflect the regional economic, social and resource conditions to a certain extent.

(1) Economic development level (ECO): expressed as the logarithm of per capita GDP;
(2) Science and technology expenditure (ST): expressed by the proportion of science and technology expenditure in fiscal expenditure;
(3) Human capital (HUM): expressed as the proportion of the number of students in Colleges and universities to the total population;
(4) Foreign direct investment (FDI): expressed as the proportion of foreign direct investment in GDP;
(5) Financial development (FIN): expressed as the proportion of the balance of deposits and loans of financial institutions in GDP (Ma & Wang, 2022);
(6) Environmental regulation (ER): expressed by the proportion of the frequency of environmental protection words in the work report of prefecture level municipal government.

The descriptive statistics of all variables are shown in Table 1.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>GI</td>
<td>2,538</td>
<td>5.1888</td>
<td>1.652</td>
<td>0.693</td>
<td>10.454</td>
</tr>
<tr>
<td>DE</td>
<td>2,538</td>
<td>0.0000</td>
<td>1.215</td>
<td>-1.476</td>
<td>12.541</td>
</tr>
<tr>
<td>ECO</td>
<td>2,538</td>
<td>7.3690</td>
<td>0.917</td>
<td>4.903</td>
<td>10.550</td>
</tr>
<tr>
<td>ST</td>
<td>2,538</td>
<td>1.6490</td>
<td>1.666</td>
<td>0.067</td>
<td>20.684</td>
</tr>
<tr>
<td>HUM</td>
<td>2,538</td>
<td>1.8940</td>
<td>2.445</td>
<td>0.004</td>
<td>13.112</td>
</tr>
<tr>
<td>FDI</td>
<td>2,538</td>
<td>1.6670</td>
<td>1.774</td>
<td>0.000</td>
<td>19.937</td>
</tr>
<tr>
<td>FIN</td>
<td>2,538</td>
<td>2.8510</td>
<td>2.272</td>
<td>0.588</td>
<td>38.237</td>
</tr>
<tr>
<td>ER</td>
<td>2,538</td>
<td>0.2510</td>
<td>0.145</td>
<td>0.000</td>
<td>1.239</td>
</tr>
</tbody>
</table>

2.3 Data Source and Processing

Considering the availability of data, we finally used the data of 282 cities in China from 2011 to 2019 for empirical research. In particular, the digital inclusive finance index comes from the digital inclusive finance index system and index compilation, the green patent application data comes from Chinese Research Data Services (CNRDS) Platform, and other data come from
China Urban Statistics Yearbook, China Environmental Statistics Yearbook and China Information Industry Yearbook. Some missing data were obtained by linear interpolation.

3 Results & Discussion

3.1 Regression Analysis

First, we tested the global Moran index of digital economy and green innovation from 2011 to 2019, and the result is significantly positive, which indicating that between the two has a strong spatial correlation. And then, Hausman test, LR test, LM test and Wald test are all significant at the level of at least 5%, which indicates that choosing the spatial Dubin model (SDM) with spatiotemporal double fixed effects is appropriate.

Table 2: Estimation results of spatial effects

<table>
<thead>
<tr>
<th>Variables</th>
<th>W1</th>
<th>W2</th>
<th>W3</th>
</tr>
</thead>
<tbody>
<tr>
<td>DE</td>
<td>0.102***</td>
<td>0.114***</td>
<td>0.090**</td>
</tr>
<tr>
<td></td>
<td>(0.029)</td>
<td>(0.029)</td>
<td>(0.029)</td>
</tr>
<tr>
<td>WxDE</td>
<td>0.253***</td>
<td>0.252***</td>
<td>0.100***</td>
</tr>
<tr>
<td></td>
<td>(0.115)</td>
<td>(0.070)</td>
<td>(0.046)</td>
</tr>
<tr>
<td>Direct effect</td>
<td>0.110***</td>
<td>0.117***</td>
<td>0.098***</td>
</tr>
<tr>
<td></td>
<td>(0.029)</td>
<td>(0.029)</td>
<td>(0.029)</td>
</tr>
<tr>
<td>Indirect effect</td>
<td>0.315***</td>
<td>0.279***</td>
<td>0.151**</td>
</tr>
<tr>
<td></td>
<td>(0.067)</td>
<td>(0.071)</td>
<td>(0.052)</td>
</tr>
<tr>
<td>Total effect</td>
<td>0.425***</td>
<td>0.397***</td>
<td>0.249***</td>
</tr>
<tr>
<td></td>
<td>(0.070)</td>
<td>(0.073)</td>
<td>(0.056)</td>
</tr>
<tr>
<td>Control variables</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Observations</td>
<td>2538</td>
<td>2538</td>
<td>2538</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.7658</td>
<td>0.7515</td>
<td>0.5404</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-649.3888</td>
<td>-661.1476</td>
<td>-612.0804</td>
</tr>
</tbody>
</table>

Table 2 reports the SDM results under the economic geography nested matrix (W1), economic geography matrix (W2) and adjacency matrix (W3). In the first column of Table 2, the digital economy is significantly positively correlated with green innovation (α=0.102, p=0.000), and has significant spatial spillover effect (α=0.253, p=0.000). After further analysis of the test results, we find that the direct effect of digital economy on green innovation is significantly positive (α=0.110, p=0.000), the indirect effect is significantly positive (α=0.315, p=0.000), the total effect is also significantly positive (α=0.425, p=0.000). The results supports the research hypothesis, and the results of W2 and W3 also strengthen this conclusion.

3.2 Regional Heterogeneity Analysis

Due to digital economy in different regional resources and development stages may have differences, the impact on green innovation may also be heterogeneous. Referring to the division of the three economic zones, we divide 282 cities into three parts. Table 3 shows the SDM results of the eastern, central and western regions under W1.
Table 3 Urban scale heterogeneity analysis

<table>
<thead>
<tr>
<th>Variables</th>
<th>Eastern region</th>
<th>Central region</th>
<th>Western region</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DE</td>
<td>0.001 (0.047)</td>
<td>0.140** (0.054)</td>
<td>0.073 (0.054)</td>
</tr>
<tr>
<td>WxDE</td>
<td>0.139 (0.097)</td>
<td>0.670*** (0.134)</td>
<td>0.014 (0.104)</td>
</tr>
<tr>
<td>Direct effect</td>
<td>0.010 (0.049)</td>
<td>0.156** (0.055)</td>
<td>0.075 (0.055)</td>
</tr>
<tr>
<td>Indirect effect</td>
<td>0.173 (0.117)</td>
<td>0.787*** (0.146)</td>
<td>0.005 (0.094)</td>
</tr>
<tr>
<td>Total effect</td>
<td>0.183 (0.133)</td>
<td>0.942*** (0.148)</td>
<td>0.080 (0.105)</td>
</tr>
<tr>
<td>Control variables</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Observations</td>
<td>1017</td>
<td>972</td>
<td>549</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.532</td>
<td>0.784</td>
<td>0.041</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-191.165</td>
<td>-261.849</td>
<td>-146.291</td>
</tr>
</tbody>
</table>

We can clearly know that the digital economy in central region cities has significantly promoted the development of green innovation at the level of 5%, but it is not significant in the other regions. The possible reason is that the central region has better digested and absorbed the cash knowledge and technology of the eastern region, seized the digital dividend generated by the digital economy, and thus showed a stronger willingness to green innovation.

3.3 Urban Scale Heterogeneity Analysis

Due to large differences in urban population, referring to the practices of He et al. (2020)[1], divide into small-medium sized cities and large cities according to whether the urban population is less than 5 million. Table 4 shows the analysis of urban scale heterogeneity.

Table 4 Urban scale heterogeneity analysis

<table>
<thead>
<tr>
<th>Variables</th>
<th>Big</th>
<th>Small-medium</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DE</td>
<td>0.156*** (0.046)</td>
<td>0.070** (0.036)</td>
</tr>
<tr>
<td>WxDE</td>
<td>0.304** (0.105)</td>
<td>0.202** (0.070)</td>
</tr>
<tr>
<td>Direct effect</td>
<td>0.161*** (0.047)</td>
<td>0.076** (0.037)</td>
</tr>
<tr>
<td>Indirect effect</td>
<td>0.322** (0.102)</td>
<td>0.239*** (0.075)</td>
</tr>
<tr>
<td>Total effect</td>
<td>0.483*** (0.102)</td>
<td>0.315*** (0.079)</td>
</tr>
<tr>
<td>Control variables</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Observations</td>
<td>819</td>
<td>1719</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.7662</td>
<td>0.6883</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-85.8086</td>
<td>-533.9603</td>
</tr>
</tbody>
</table>
The impact in large cities is positive and significant ($\alpha=0.156$, $p=0.000$), while the impact on small-medium sized cities is also positive and significant ($\alpha=0.070$, $p=0.007$). However, the impact of big cities is stronger in terms of value and significance. This may because big cities have a good foundation in R&D foundation, digital technology and industrial agglomeration, forming the characteristics of coordinated evolution of digital economy and green innovation.

### 3.4 Robustness and endogenous text

#### 3.4.1 Robustness Check

We conducted a series of robustness tests to further enhance the robustness of the results. Firstly, we use the authorized number of urban green innovation patents as the new explained variable ($\alpha=0.118$, $p=0.000$). The SDM results based on W1 are shown in column (1) of Table 5. It is found that the conclusion is still supported. Secondly, due to the low administrative levels in terms of political resources and innovation ability, we exclude cities with high administrative levels (municipalities and provincial capitals), and only 247 general cities are retained as samples for regression. The SDM results under W1 are shown in column (2) of Table 5. The research conclusions and are still robust ($\alpha=0.083$, $p=0.000$).

<table>
<thead>
<tr>
<th>Variables</th>
<th>Green patents authorized number</th>
<th>Ordinary prefecture level city</th>
</tr>
</thead>
<tbody>
<tr>
<td>DE</td>
<td>$0.118^{***}$ (0.028)</td>
<td>$0.083^{**}$ (0.032)</td>
</tr>
<tr>
<td>WxDE</td>
<td>$0.143^{**}$ (0.060)</td>
<td>$0.175^{**}$ (0.065)</td>
</tr>
<tr>
<td>Direct effect</td>
<td>$0.123^{***}$ (0.028)</td>
<td>$0.089^{**}$ (0.033)</td>
</tr>
<tr>
<td>Indirect effect</td>
<td>$0.184^{**}$ (0.065)</td>
<td>$0.213^{**}$ (0.071)</td>
</tr>
<tr>
<td>Total effect</td>
<td>$0.307^{***}$ (0.067)</td>
<td>$0.302^{***}$ (0.075)</td>
</tr>
<tr>
<td>Control variables</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Observations</td>
<td>2538</td>
<td>2223</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.698</td>
<td>0.714</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-591.964</td>
<td>-633.981</td>
</tr>
</tbody>
</table>

#### 3.4.2 Endogenous Text

Since the two-way causality will affect the accuracy of the research results, we select the digital economic variables lag for one period (GI1) as the instrumental variable, and use the two-stage least square method for regression. The endogenous text results are shown in Table 6. After considering endogeneity, the positive impact of digital economy on green innovation is still significant. Meanwhile, Kleibergen-Paap rk LM statistics p value is 0.000 ($\alpha = 442.156$), which indicates that the instrumental variables can confirm the research hypothesis; in the test of weak identification of instrumental variables, the Kleibergen-Paap rk Wald F statistic is greater than the critical value at the 10% level of Stock-Yogo weak identification test. Overall, it is reasonable to choose GI1 as the tool variable of digital economy.
### Table 6 Robustness check

<table>
<thead>
<tr>
<th>Variables</th>
<th>2sls</th>
</tr>
</thead>
<tbody>
<tr>
<td>GI1</td>
<td></td>
</tr>
<tr>
<td>DE</td>
<td>2.284**</td>
</tr>
<tr>
<td></td>
<td>(0.058)</td>
</tr>
<tr>
<td>Control variables</td>
<td>YES</td>
</tr>
<tr>
<td>Kleibergen-Paap rk LM statistic</td>
<td>442.156***</td>
</tr>
<tr>
<td>Kleibergen-Paap rk Wald F statistic</td>
<td>4311.507</td>
</tr>
<tr>
<td>Observations</td>
<td>2256</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.782</td>
</tr>
</tbody>
</table>

### 4 Conclusion

Based on the data of 282 Chinese cities in 9 years, this study verifies the impact mechanism between digital economy and green innovation by using SDM. The results show that the digital economy can promote the development of green innovation through space spillover effect. Further research found that the digital economy of center regions and large cities can promote the development of green innovation.

This study has the following policy significance. (1) Clarify the important role of digital economy in promoting the development of urban green innovation. Further promote the full coverage of 5G, big data and other digital infrastructure as soon as possible to lay a solid material foundation for the release of digital dividends. At the same time, increase the investment in digital industry, build a platform for digital technology and industry integration, and enable the development of industries from multiple angles and in all directions. (2) Fully consider the spatial spillover effect of digital economy, and implement the regional difference strategy. There are differences in the promotion role of digital economy in cities of different sizes and different regions, use digital technology to build a sharing mechanism of technology, information, talent and other resource elements between different regions, and make full use of the "diffusion effect" of big cities to drive the development of surrounding cities.

However, our research still has some limitations. (1) Although the study surveyed most cities in China, due to the availability of data, we have no way to analyze all cities in China. In addition, due to the restrictions of COVID-19, our research period is up to 2019. Future research can further expand the research sample and interval. (2) The connotation of digital economy may change with the passage of time. Therefore, future research can further expand the comprehensive index of digital economy. (3) The study only takes cities in China as an example. If cities in other countries are compared, new findings may be made.

### REFERENCES


Research on E-Commerce Trading System Market-Oriented Transformation Which Based the Value Co-Creation Theory

Ming Shi\textsuperscript{1,a}, Ming Zhang\textsuperscript{2,b}, Yu Duan\textsuperscript{2,c}

shihongming@sgdt.sgcc.com.cn\textsuperscript{a}, zhangming@sgdt.sgcc.com.cn\textsuperscript{b}, duanyu@sgdt.sgcc.com.cn\textsuperscript{c}

State Grid Digital Technology Holding Co., Ltd, Xianglong Business Building, 311 Xicheng District, Beijing, China\textsuperscript{1}

State Grid Zhilian E-commerce Co., Ltd, Maoye Building, 78 Haihe East Road, Hebei District, Tianjin, China\textsuperscript{2}

Abstract: With the further acceleration of the digital transformation of energy state-owned enterprises, e-commerce platform, as an important starting point of the digital transformation of energy state-owned enterprises, is in urgent need of a comprehensive upgrade through the market-oriented transformation of the trading system. Based energy marketization of electric business platform trading system transformation necessity in state-owned enterprises, focusing on electric business platform supplier, Taipei, the value of the three major trading main body to create purchasing customers, play a leading role of typical cases and reference market head electric business platform trading mode innovation, market-oriented transition route of soe electric business platform, clear electric business platform trading pattern innovation focus, To innovate and build strategic paths, drive industrial units of provincial management to take value co-creation as the vision, and promote the digital, intelligent and networked development of e-commerce trading system.

Keywords: Energy State-Owned Enterprises, Value Co-Creation, Trading System, Transition to Market.

1 INTRODUCTION

In March 2021, China issued 《the 14th Five-Year Plan for National Economic and Social Development of the People’s Republic of China and the Outline of Long-term Goals to 2035》, which clearly put forward the promotion of digital transformation. As an important scene of the digital transformation of state-owned energy enterprises, e-commerce platform's transaction mode is the core system to promote the integration of digital economy and real economy. Therefore, the market-oriented transformation of the trading system of the e-commerce platform of energy state-owned enterprises is crucial, which can effectively solve the pain points of the traditional industries in commodity trading, warehousing and logistics, processing and transportation, and promote the integrated development of capital flow, information flow and logistics. As the backbone state-owned enterprise in the energy industry, State Grid actively plays a leading role and takes the lead in the marketization transformation of the trading system.
2 NECESSITY OF MARKETING TRANSFORMATION

With the integration of energy revolution and digital revolution, it has become a general trend to cultivate a new digital ecology of energy through digital innovation. As an important starting point for the digital transformation of state-owned energy enterprises, the market-oriented transformation of e-commerce trading system is the only way to comply with the national "digital transformation" strategy, promote the integration of the digital economy and the real economy in the energy industry, and support the e-commerce of enterprise procurement transactions.

2.1 Key Measures in the National Digital Transformation Strategy

Under the guidance of the "Notice on Accelerating the Digital Transformation of State-owned Enterprises" issued by SASAC, the digitalization, networking and intelligence of energy state-owned enterprises are accelerated. As an important interaction scene of information flow, capital flow and logistics, e-commerce platform is the main front of digital transformation. At the same time, the Notice on Accelerating the Development of Industrial Internet further clarifies the necessity of turning the trading platform of state-owned energy enterprises into a trading platform, and promotes the iteration of the procurement trading platform from an information platform to a trading platform. It is imperative to enhance the trading service capability of the platform.

2.2 The Important Focus of Enabling Industry Chain Development

The energy industry and industrial products industry have the characteristics of upstream and downstream dispersion and large gap in individual scale, which brings the value space of deep link for platform transactions. The innovation of the trading system of the energy state-owned enterprises focusing on the e-commerce platform can further strengthen the cooperation between the platform and upstream suppliers, strengthen the downstream regional network layout, actively develop small and medium-sized enterprise customers, and leverage the business growth.

2.3 Important Way to Support Enterprise Procurement E-Commerce

Under the normal epidemic situation, purchasers will be more inclined to use online communication and procurement\(^4\), and enterprises' demand for e-commerce procurement continues to rise. As an important system for the e-commerce platform of state-owned energy enterprises to achieve procurement transactions, the trading system can provide basic support for the innovation and upgrading of the e-commerce platform through the integrated application of cloud computing services, big data and other emerging digital technologies, and further promote the deep integration of online and offline procurement services.

3 THE DRIVING FORCE OF VALUE CO-CREATION IN E-COMMERCE PLATFORM DEVELOPMENT

Value co-creation is an important guide for resource allocation in the new economic era, and an important mode to further promote the interaction of information flow, capital flow and logistics.
In the face of this, the provincial management industry e-commerce platform can be driven by value co-creation, deeply explore the interactive value of trading subjects, give play to the value of the platform, and realize the comprehensive innovation and upgrading of the trading system mode.

3.1 Value Co-Creation Connotation

Value co-creation is a joint process in which the platform and customers create value through interaction\(^2\). It needs to realize the integration of knowledge-based resources or open resources. Its important feature is that it takes place in the context of the ongoing communication between suppliers and purchasing customers. When the e-commerce platform, purchasing customers and suppliers interact with each other to transfer value, the three will form a system to capture and meet the demand, so as to jointly create value.

3.2 The Driving Force of Value Co-Creation in the Market-Oriented Transformation of The Platform

Through value to create a vision, can further mining province tube industry electric business platform suppliers, procurement customers, ping island's three big trading subject between value creating, promote the expansion of platform trading service boundaries, with typical electric business platform in the value creating under the perspective of system innovation as the guide, for the province tube industry electricity market trading platform trading system to provide support.

3.2.1 Value Co-Creation is a New Force to Speed the Development of Platform.

The interaction of the provincial management industry e-commerce platform is embodied as follows: centering on the needs of downstream customers, it strives to form a community of interests between downstream customers and upstream suppliers. To achieve its purpose, value creating become province tube industry electric business platform for further innovation platform system mode, expand the platform service boundary important claims\(^3\), by providing products, services and solutions to coordinate the supply chain of each link, to achieve the sharing of information, capital and logistics form business ecosystem, contribute to various stakeholders coevolution, Achieve the goal of the development and growth of the entire business ecosystem.

3.2.2 Market Cases as New Force to Guide the Development of Platform.

At present, studies on value co-creation are mainly focused on purchasing customers and suppliers, and service leading logic and purchasing customer experience logic respectively emphasize the leading role of suppliers and purchasing customers in value co-creation (Jian, Xiao, 2015). As an important participant in the value co-creation mode of e-commerce platform, the provincial management industrial e-commerce platform can not only eliminate consumption barriers, but also build the interaction channel between purchasing customers and suppliers, and ultimately promote the customized production of suppliers. Therefore, the demonstration and leading role of typical e-commerce platforms should be further given play, and the exploration and practice of mainstream e-commerce platforms under the perspective of compact value co-creation should be summarized, so as to provide reference for the comprehensive upgrade of the trading system of provincial e-commerce platforms.
4 BASED ON THE PERSPECTIVE OF VALUE CO-CREATION, THE ANALYSIS OF TYPICAL DEMONSTRATION CASES

Create a less focus on the platform based on the value of the status quo, the province tube industry unit to market-oriented industrial electricity head platform, market-oriented transformation completed state electric business platform for the demonstration of typical cases and analysis of its value to create a trading system under the perspective to explore, for the province tube industry electricity trading platform trading patterns of innovation.

4.1 Case on the Innovation of Market E-Commerce Platform Trading System

Focusing on the exploration of the trading system mode of Zhenkunxing, Jingdong Industrial Products and Alibaba Industrial Products Station, the three leading platforms of market-oriented industrial products e-commerce in the perspective of value co-creation, it provides guidance and reference for the innovation and upgrading of the trading system of provincial management industrial e-commerce platforms.

4.1.1 Zhenkunhang Supermarket.

Based on the perspective of value co-creation, Zhenkunhang finds the pain points of MRO industry, realizes the platform transaction mode innovation with "perfect customer experience" as the core, and enhances the user's stickiness on the platform with intelligent services.

Intelligent solutions with suppliers and purchasing customers realize the value of creating. Relying on the one-stop online procurement platform, through product optimization, online intelligent warehousing management, information technology and data capabilities, technical services and other service solutions, to provide data resource support for suppliers, to provide intelligent solutions for purchasing customers.

Strong organizational support make online supply chain integration possible. Design risk control system, independent logistics system, regional service center and talent team four support systems. Through the combination of e-commerce and intelligent hardware, Internet to the Internet of things application, to solve the customer's one-stop procurement needs, on-site technical services and personalized logistics needs, to achieve the value of the whole chain of the supply chain partners.

4.1.2 Jingdong Industrial Products.

Jingdong create financial solutions and activate supply chain service value. Based on years of accumulation in the field of supply chain management, JD provides financial solutions for upstream suppliers, core enterprises and downstream vendors to achieve complete supply chain coverage, that is, to provide factoring services for upstream suppliers. For core enterprises, there is payment settlement (Ding, Zhan, 2021); For downstream sellers and end customers, there are a variety of credit products; For the warehouse transportation links, including chattel mortgage and freight factoring, the multi-dimensional financial services are used to achieve the aggregate management of the capital flow in the supply chain, forming a mutually beneficial and sustainable development of the supply chain financial ecosystem.
4.1.3 Alibaba Industrial Product Station.

From the perspective of value co-creation, Alibaba Industrial Product Station uses big data to segment customers and provide data services to realize its own value.

**Segment procurement customers, pass "experience value"**. Alibaba Industrial Products Website collects purchasing customers' data on the e-commerce platform, and analyzes each user's search content, product browsing details, length of stay on the page, compared products and last purchased products by establishing an individual identification model. Furthermore, differentiated precision marketing can be carried out by identifying the needs and preferences of purchasing customers.

**Provide data services, pass to suppliers "demand information value"**. Based on the massive data accumulated by Alibaba's industrial product website, through multi-dimensional analysis of transaction time, commodity price, transaction quantity, enterprise type and other information, the "demand information value" is conveyed to suppliers (Yan, He, 2017), and the business advisory board is created for suppliers to provide business analysis, market conditions, special tools and other services. Assist suppliers in market positioning, sales planning, investment planning, production planning, capital demand planning and other activities, effectively allocate capacity and reduce inventory.

4.2 Case on the Market-Oriented Transformation of State-Owned Enterprises' E-Commerce Platforms

Focusing on Datang Electronic trading platform and Epaike Industrial products e-commerce platform, two state-owned industrial products e-commerce platforms in market transformation, this paper analyzes their successful advanced experience in market transformation from the perspective of value co-creation, so as to provide reference for the market-oriented transformation of the provincial management industrial e-commerce platform trading system.

4.2.1 Datang Electronic Trading Platform

Datang e-commerce platform focuses on the value co-creation between suppliers and platforms, as well as between purchasing customers and platforms, and standardizing the trading behaviors of various procurement subjects through the procurement management system.

**To optimize supplier evaluation system, reconstructing value idea of cooperation**. Through the management platform, Datang e-commerce platform comprehensively establishes supplier files, combines daily management with annual assessment, and strengthens the management of supplier performance evaluation and bad behavior. According to the evaluation results of suppliers, suppliers shall be managed by classification, forming three types of suppliers: registration, access and strategy, and building a first-class supplier system featuring high efficiency, high quality, mutual trust, cooperation and win-win results.

**Improve purchasing management system, implementation of accurate monitoring**. While strengthening the construction of the procurement supply chain system, Datang e-commerce platform constantly improves the procurement management system, builds the "big supervision" system, builds the procurement supervision system, implements the procurement supervision through the combination of online and offline, and escorts the standardized operation of procurement at all levels.
4.2.2 EPEC Industrial platform

EPEC Industrial products e-commerce platform further integrates the digital intelligence supply chain by focusing on all transaction subjects in the supply chain and multi-dimensional creation of the value of EPEC platform, and at the same time realizes the innovation and upgrade of the platform trading mode.

Focused on the subject value of supply chain to create, to realize the end-to-end chain cover. Strengthen the internal and external coordination of the supply chain, promote the real-time, transparent, interconnected and traceable links of the supply chain including demand planning, bidding and procurement, manufacturing management, quality inspection, warehousing and distribution, logistics and delivery through digital technology, and improve the flexibility of the supply chain.

Further play electric business platform value, promote the closed-loop management of intelligent operation. Show the business situation in real time through the operation Kanban, perceive, monitor, discover and solve the problems from suppliers, manufacturing, products, transportation and customers in 360 degrees, and truly realize the improvement of efficiency, create value and form a new digital supply chain.

4.3 Summary of Typical Paths

Based on value creating perspective, kun line, jingdong industrial goods, alibaba industrial stand three market head platform trading pattern innovation to upgrade, and datang electronic trading platform, easy to send the guest industrial advanced experience of the two energy platform market-oriented transformation of state-owned enterprises, further defined the province tube industrial electric business platform trading pattern innovation point.

4.3.1 Focus on the Behavior of Trading Entities, Empowered by Emerging Digital Technologies.

The three major trading subjects of e-commerce platform are suppliers, purchasing customers and platforms. Through the multi-dimensional application of five digital technologies of "cloud, large, material, mobile and intelligence", behavioral data of trading subjects can be further collected and value co-creation among platform, purchasing customers and suppliers can be continuously explored.

4.3.2 Build a Business Ecosystem Driven by Service Model Innovation.

Connect upstream and downstream enterprises through the platform mode, connect the supply and demand of high-quality products, solve the industry pain points, reconstruct the supply chain, and bring connection dividends for all links of the supply chain; Platform in order to increase the viscosity and improve the comprehensive competitiveness platform, realizes the integrated service providers to shift, platform enterprises need through data services provide decision-making basis for the production and sale of the upstream enterprise and downstream of the intelligent by providing solutions to help customers improve the efficiency of procurement and management, so as to improve the structure of the whole supply chain, the formation of business ecosystem, Facilitate the co-evolution of various stakeholders.
4.3.3 Quantify the Value Dimension of the Platform under the Guidance of Building a Value Diagnosis System.

In the perspective of value co-creation, the value creation of purchasing customers and the scope and depth of the operation of the e-commerce platform are the core control points for the continued development of the overall competitiveness of the platform, which will provide important guidance for the continuous innovation and upgrading of the platform trading mode in the future. Due to differences of platform development orientation, evaluation value to create dimension index is difficult to standardize, therefore, should be according to the different links of electric business platform service interaction, degree of purchasing the diversification of customer needs, platform, resources integration and synergy mode, value to create a platform for the operation mechanism, internal output value of the actual effect, etc., every business sectors build platform value diagnosis model, Quantification of platform value enhancement degree to enhance market value.

5 E-COMMERCE PLATFORM TRADING SYSTEM MARKET TRANSFORMATION STRATEGY PATH

Province electricity pipe industry trading platform as a national power grid in the digital transformation, the number of supply chain to build, and the important fulcrum of purchasing electricity, to create a platform for value, innovation make target three big pillar + + security level, build a trading system innovation path can be born, realize marketization of trading system transformation and upgrading.

5.1 Target Layer: Create a Whole-Domain E-Commerce Service Trading System with Value Co-Creation as the Core

In order to effectively ensure the realizability of goals and the landing of strategic paths, three principles will be followed when setting goals for the provincial management of industrial e-commerce trading system, namely, basing on the existing development goals of enterprises, realizing the value co-creation of transaction subjects, and expanding the trading service boundary of the platform.

5.1.1 Based on the Existing Development Goals of the Enterprise.

As an important carrier of the industrial product e-commerce platform of the State Grid, the provincial Management industrial e-commerce trading platform takes "one body and four wings" as the strategic layout, and aims to build an internationally leading energy Internet with Chinese characteristics. At the same time, the existing services of the provincial management industry e-commerce trading platform are mainly focused on the matchmaking of deals, and the initiative ability of upstream and downstream services needs to be further enhanced.

5.1.2 To Realize the Value Co-Creation of the Transaction Subject.

Focusing on the trading behaviors of suppliers, platforms and purchasing customers, on the basis of behavioral data collection, new digital technologies such as big data and artificial intelligence are used to deeply explore the demands of existing trading behaviors on platform capabilities and platform services, and further optimize the platform trading system.
5.1.3 Expand the Trading Service Boundary of the Platform.

To build a digitalized and intelligent online and offline full business chain and all-field e-commerce service trading system, and further expand the platform's trading service boundary while improving the ecological chain's all-field operation service capability.

5.2 Pillar Layer: Subject Supervision + Service Expansion + Value Diagnosis

From the perspective of value co-creation, transaction subject supervision, service model expansion and platform value diagnosis of the provincial management of industrial e-commerce trading platform are taken as the pillar layer to form an iterative update closed loop of value co-creation.

5.2.1 Implement Whole-Process Supervision of Trading Subjects with Digital Technology.

In order to collect and analyze the behavior data of suppliers, purchasing customers and platforms in purchasing transactions, the provincial management industrial e-commerce platform further integrates digital technologies such as the Internet of Things, blockchain, big data and artificial intelligence to realize the supervision of the whole process of transaction subjects.

Suppliers: deepen the digital technology application, optimize the whole life cycle of supplier behavior regulation. Supervise the whole process of production, shelving, bid response and contract performance at the contract execution stage of suppliers by placing intelligent sensors, dynamic video surveillance and other means. Conduct equipment supervision, supervision and inspection in the production and manufacturing links of suppliers, carry out actual quality supervision and service guarantee supervision in the suppliers' standard compliance links, and carry out material delivery, acceptance, installation, use and performance evaluation in the suppliers' performance links, so as to realize the evaluation and assessment of the whole life cycle.

Purchasing customers: relying on the platform of information transparency, implement purchasing customers whole process behavior regulation. In the stage of customer demand submission, set up a clear demand submission system, online record material product parameters, service product effects and other requirements, and a review team composed of industry experts and purchasing leaders, online approval and confirmation of purchasing customers' procurement information such as purchasing plan, budget limit and autonomy. In the stage of evaluation and performance of the contract, actively urge the purchasing customers to evaluate the products and services that have completed the performance of the contract in the form of online reminder, so as to realize the responsibilities and norms of conduct of the whole subject of the procurement process.

Flat Taipei: according to the historical transaction data, fine electric business platform trading behavior regulation. Based on the historical purchase and sales data of the platform and combined with the current market situation, the change trend of purchase price, supplier storage and logistics price is analyzed to ensure that the transaction price is within the normal floating range and prevent suppliers from hoarding goods and raising prices. At the same time, the procurement frequency and submission times of procurement requirements of purchasers are monitored, and abnormal purchasing behaviors of purchasers are warned to effectively
regulate procurement transactions on the platform.

5.2.2 The Expansion of Platform Service Mode with Comprehensive Solutions

Basic services: intelligent procurement services + procurement solutions. Firstly, smart purchasing service. Statistics based on the historical data of enterprise procurement, the company yearly/quarterly/monthly purchasing directory, through the provincial electricity pipe industry big data intelligent analysis of trading platform, to provide enterprise procurement intelligence comparison, smart reminder, intelligent early warning, AI intelligent assistant, suppliers, intelligent risk control and so on many AI intelligent, drive enterprise procurement business each link intellectualization. Secondly, procurement solutions. Based on the procurement needs of enterprises, control the key links in the process of requisition, ordering, arrival, warehousing, billing and settlement, optimize the purchase price and quantity, track the implementation of the procurement progress, reduce the procurement cost, avoid the procurement risk, ensure the supply of materials, and assist enterprises in the overall management of the procurement process.

Value-added services: the financial guarantee services + energy solutions. Firstly, financial guarantee service. According to the results of big data analysis, the credit and financing limit of the enterprise is judged to ensure that the financing and factoring amount of the platform is in line with the situation of the enterprise and the capital is controllable. Further, the provincial management of industrial e-commerce trading platform can cooperate with banks and financial guarantee institutions to build a supply chain financial service platform, transform the platform data into the credit line recognized by banks and financial guarantee institutions, and then complete independent approval and loan issuance. At the same time, relying on the huge customer resources of energy state-owned enterprises, we reach cooperation with commercial factoring companies and apply for commercial factoring services with the guarantee of business documents. Secondly, energy solutions. As an e-commerce platform for industrial products in the national grid system, the Provincial Management Industry e-commerce trading platform can rely on the frontier interpretation of the energy and power industry by the State Grid to provide value-added services such as power and energy industry analysis, power saving transformation, power consumption data analysis and customized comprehensive energy solutions.

5.2.3 Realize the Value Diagnosis of E-Commerce Platform with Multiple Evaluation Dimensions.

In the context of the dynamic and complex Internet era, in order to further realize the "1+1>2"overall value added effect, provincial management of industrial e-commerce trading platform to establish a value diagnosis system, screening supply chain synergy value, platform trading value, market expansion value as an important measurement dimension, to provide diagnostic suggestions for platform value enhancement.

Collaborative supply chain value: save electricity pipe industry trading platform of the trading main body in the supply chain upstream and downstream business activities have close correlation value, therefore, resources integration and service interaction value are two important dimensions of supply chain collaboration. Firstly, value of resource integration: expanding business channels + acquiring actual resources. Resource integration is an important index to realize professional and diversified business activities of upstream suppliers
on the platform and to measure the actual effect of supply chain collaborative value output. Therefore, the provincial management of industrial e-commerce trading platform can use the resource integration value as a diagnostic node to evaluate the practical effects of expanding the types of business channels and enriching access to resources in the supply chain synergy value function block of the provincial management of industrial e-commerce trading platform, guide and constantly update various value-added services of the research and development platform. To realize the collaborative operation of each participant’s own advantageous resources and the high-quality resources of other participants, and promote the value creation of the provincial management industrial e-commerce trading platform.

**Secondly, service interaction value:**

**service type + service interaction promotion.** The platform, upstream suppliers, downstream procurement customer service interaction behavior, interaction dimension, interaction frequency and other indicators are the direct embodiment of the evaluation platform supply chain collaboration value. Therefore, the service interaction value can be used as the diagnostic node to evaluate the actual effects of the improvement of service mode and the promotion of service interaction process of the provincial management industrial e-commerce trading platform, so as to bring more high-quality products and services to the platform and further meet the differentiated demand experience of customers.

**Platform trading value:** save electricity pipe industry trading platform as a national power grid in important exploration of industrial electric business platform, promote the platform transaction value is to speed up purchasing online, accelerate energy industrial electric business platform for the development of important measures in state-owned enterprises, and product innovation, the economic benefit is an important dimension of measuring platform trade value.

**Firstly, product innovation value:** product service quality + product performance improvement. Through product creative planning, business process reengineering, process improvement and other ways, to constantly improve the product performance of the platform's core business links is a key move to ensure the quality of the platform's products and services. Therefore, product innovation value can be used as a diagnostic node to evaluate the practical effects of enhancing product service quality and improving product performance in the platform transaction value function block of the provincial management industrial e-commerce platform.

**Secondly, economic value:** Profitability + cash flow. The e-commerce platform can adopt the operation mode of the whole network channel to improve the efficiency of operation management, obtain resources more efficiently through the value co-creation activities of trading subjects in the platform, and continuously improve the business link of the value co-creation process of the platform through effective information feedback and business data, so as to realize the next stage of platform value added. Therefore, the economic benefit value can be used as the diagnostic node to evaluate the practical effects of the provincial management industrial e-commerce platform in the trading value function block of the platform, such as improving profitability and accelerating capital turnover.

**Market value:** the province tube industry electric business platform is focused on the key business activities of the industrial products of the Internet, according to the platform operation mode to integrate business processes, and to create a link with the value of associated nodes matching, improve the overall competitiveness. At the same time, the platform should be based on the perspective of value co-creation, realize multi-party cooperation, establish an open, open, inclusive and fair value co-creation operation system, and maximize the economic benefits of the platform. Therefore, the market expansion value can be used as the diagnostic node to
evaluate the practical effects of improving market competitiveness and expanding the operation scale of the platform transaction value function block of the provincial regulated industrial e-commerce platform.

5.3 Guarantee Layer: Build Data System and System Functions

Province tube industry electric business platform to build the data system and development system module as the support, by digital technology such as large data calculation, artificial intelligence, provide for the customer to choose, logistics and financial and other one-stop service procurement solutions, implementation in whole life cycle of electric business platform for the trading main body behavior management digitization, automation, intelligent, Thus, the service quality of the platform trading mode can be improved, as shown in the figure below.

5.3.1 Multi-Source Data Input

In order to avoid over-reliance on a single data source, which may cause invalid analysis, multi-source data such as platform transaction data, industry professional database, supplier and purchasing customer referral data can be introduced, and cross-verification mechanism between different data sources can be established.

5.3.2 Iterative Data Model

In order to realize the value of data analysis, e-commerce platforms need to continuously iterate data models, especially data statistics and comparison of the same category, category price calculation, procurement demand forecast, etc.

5.3.3 System Module Function

Through big data technology, platform service functions such as standardization of commodity management, intelligent product selection and procurement plan formulation, intelligent contract performance and delivery, and intelligent evaluation can be realized.

6 CONCLUSION

This paper makes an in-depth interpretation of the necessity of market-oriented transformation of the trading system of the current state-owned energy e-commerce platform, focuses on the innovation and upgrading of the trading mode of the e-commerce platform from the perspective of value co-creation, summarizes the innovation of the trading mode of the typical e-commerce platform based on the value co-creation theory, and builds a three-level strategic path of "goal + support + guarantee". Implement the digital, intelligent and networked development of the trading system of the provincial management industry e-commerce trading platform. Follow-up, the province tube industry electricity trading platform can be around existing state grid strategic path, and actual operations platform needs, continuously optimize platform trading system mode, boost the electricity province tube industry in the comprehensive upgrade of the trading platform at the same time, for other energy electric business platform the market-oriented transformation of state-owned enterprises provide reference path.
REFERENCES

Environmental Regulation, Technological Innovation and Industrial Structure Optimization

Xiaoyan Ren 1,a, Yaning Li 2,b, Xiaoqi Qin 3,c
*corresponding author: arenxiaoyan999@163.com
b-e-mail: liyaning1128@163.com
c-e-mail: 493024389@qq.com

School of Economics and Management, Xi’an University of Technology, Xi’an, China

Abstract: Based on panel data from 30 Chinese provinces between 2008 and 2019, this paper empirically analyzes the indirect effect of environmental regulation intensity on the rationalization and advancement of industrial structure in the national, eastern, central and western regions, and the moderating role of regional economic development level in the optimization of industrial structure influenced by environmental regulation intensity, using the mediating effect model and moderating effect model. The results of this study show that environmental regulation in the national, eastern and central regions affects the rationalization and advancement of industrial structure by influencing the intensity of technological innovation, but environmental regulation in the western region does not have a significant effect on the rationalization and advancement of industrial structure.

Key Words: Environmental Regulation Intensity, Technological Innovation Intensity, Industrial Structure

1. Introduction

Industrial structure issues like high energy consumption and severe pollution have grown more challenging as industrialization has progressed, making it difficult for China’s economy to grow sustainably. The rationalization and advanced nature of industrial structure are encouraged by the optimization and upgrading of industrial structure, which also helps to lower the proportion of industries with high energy consumption and high pollution, while technological innovation is the main driving force for optimizing resource allocation and promoting the optimization and upgrading of industrial structure. Exploring how environmental regulation affects industrial structure optimization under the dual pressure of environmental issues and sustainable economic development is beneficial for revealing the inner mechanism of industrial structure optimization from a theoretical perspective, as well as for advancing China’s sustainable economic development.

Some scholars believe that environmental regulation will drive up business costs more and reducing investment in technological innovation. (Zhang & Lu, 2018[1]; Gutierrez & Teshima, 2018[2]). Some scholars believe that effective environmental legislation can encourage businesses to pursue technological innovation. (Yu & Cui, 2019[3]; Shi & Zhao, 2018[4]). Some scholars believe that environmental regulation's effects on businesses' technical innovation are...
defined by cyclical shifts. (Dong & Wang, 2019[5]). In the existing research, numerous research have been done on how technology innovation affects industrial structure, but most scholars' research mainly focuses on the impact of technological innovation on industrial structure transformation and upgrading, and most of them believe that technological innovation can promote industrial structure transformation and upgrading (Bai & Zhou, 2017[6]; Xu & Liu, 2016[7]). Existing research primarily examines the influence of environmental regulation on technological innovation and the impact of technological innovation on the transformation and upgrading of industrial structures. However, comparatively little attention has been given to the effects of environmental regulation on optimizing industrial structures. And the existing literature focuses mainly on the direct impact of environmental regulation on the advancement of industrial structure, while there is a paucity of research on the indirect influence of these regulation on optimizing industrial structure.

2. Theoretical Analysis and Hypothesis Proposed

2.1 The role of technology innovation as a mediator in the optimization of industrial structure influenced by environmental regulation

In the short run, environmental regulation result in higher costs for treating pollution. Nevertheless, over time, environmental regulation are beneficial for fostering technological advancement and encouraging industrialization restructuring. (Xue, 2016[8]). This paper argues that environmental regulation are likely to spur technological innovation in enterprises. Technological innovation can accelerate the agglomeration of production factors, transform the operational and production methodologies of corporations, influence the factor structure and industrial organization, and potentially disrupt the balance of factor allocation, leading to a discordance in industrial structure. (Li & Dong, 2018[9]). And technological innovation can promote the efficient allocation of the input factors, facilitating the transition and advancement of traditional industries and fostering the growth of sustainable emerging industries. (Ren & Yang, 2020[10]). Meanwhile, technological innovation also plays a role in enhancing labor productivity and transforming the structure of societal needs, thus promoting industrial structure upgrading (Cai & Wang, 2018[11]). This paper argues that the increase of environmental regulation intensity will make the production costs of pollution-intensive industries increase, thus leading to the decrease of enterprise profits (Li et al, 2019[12]). But due to the characteristic of technological innovation such as long cycle, high risk and high investment, it makes some high pollution enterprises invest their resources in non-pollution and high-return tertiary industries. In addition, a moderate level of environmental regulation encourages enterprises to use resources effectively and enhance production efficiency by increasing the intensity of technological innovation, thus promoting the development of industrial structure from lower to higher levels.

H1: Technological innovation intensity plays a mediating role in the process of environmental regulation intensity affecting the rationalization of industrial structure/the advancement of industrial structure.
2.2 The role of technological innovation in mediating the optimization of industrial structure influenced by environmental regulation by sub-region

As environmental regulation become more stringent, enterprises tend to obtain new processes and technologies through technology innovation for the purpose of meet the government’s environmental requirements. This technology innovation will further influence the rationalization and advancement of industrial structure. Because of the relatively underdeveloped economy and the limited availability of innovation resources in the western region, enterprises’ innovation vitality and innovation motivation are insufficient, so when the government adopts strict environmental regulation policies, it tends to force enterprises to meet the government’s environmental requirements through measure such as increasing pollution control costs rather than through technical innovation, which will reduce enterprises’ technological innovation investment and inhibit their technological innovation behavior. The level of economic development and technological innovation capability in the central region are located between the eastern and western regions. So when the environmental regulation intensity increases, some enterprises may achieve government’s environmental requirements through technological innovation, but some enterprises will achieve the government’s environmental requirements by “polluting first and treating later” due to insufficient technological innovation capability.

H2a: The influence of technological innovation intensity on industrial structure has regional heterogeneity.

3. Research Design

3.1 Variable selection and measure

This paper draws on the approach of Gan et al. to measure the rationalization of industrial structure by using the Thiel index. The advancement of industrial structure is measured by the ratio of total output value of the tertiary industry to the total output value of the second industry. This study measures the intensity of environmental regulation by the amount of completed pollution control projects investment. The R&D-to-GDP ratio serves as a metric for assessing the degree of technological innovation. The region income level is used as a control variable, measured by the average wage of employee in urban units. Openness was utilized as the independent variable in this study, quantified as the proportion of foreign investment relative to GDP. To address heteroscedasticity and enhance data consistency, natural logarithms were applied to all variables.

3.2 Sample selection and data sources

The panel data from 30 Chinese provinces (excluding Hong Kong, Tibet, Macao, and Taiwan) between 2008 and 2019 has been chosen for the study. All data are obtained from the official website of the National Bureau of Statistics and further processed. To minimize the impact of outliers, all continuous variables are capped at the 1% threshold.
4. Main Model and Mediating Effect Test

4.1 Model constructing

Based on the theoretical analysis and research hypothesis, the following model is developed to test for mediating effects.

\[
\ln \text{Struct}_{it} = \alpha_0 + \alpha_1 \ln \text{Environ}_{it} + \alpha_2 \ln \text{Open}_{it} + \alpha_3 \ln \text{Income}_{it} + \epsilon_{it} \quad (1)
\]

\[
\ln \text{Tech}_{it} = \beta_0 + \beta_1 \ln \text{Environ}_{it} + \beta_2 \ln \text{Open}_{it} + \beta_3 \ln \text{Income}_{it} + \epsilon_{it} \quad (2)
\]

\[
\ln \text{Struc}_{it} = \gamma_0 + \gamma_1 \ln \text{Environ}_{it} + \gamma_2 \ln \text{Tech}_{it} + \gamma_3 \ln \text{Open}_{it} + \gamma_4 \ln \text{Income}_{it} + \epsilon_{it} \quad (3)
\]

4.2 Descriptive statistical analysis

The descriptive statistical analysis of the variables of the whole country and the eastern, central and western regions, the findings revealed notable variances in indicators between the regions, suggesting considerable regional disparities in China.

4.3 Correlation analysis

The findings indicate that, at the level of 0.05, there is a significant negative correlation between the Thiel index and the intensity of environmental regulation. The level of stringency in environmental regulation exhibits a significantly negative correlation with the progress of industrial structure, with a significance level of 0.01. The VIF values for the variables are all below 10, indicating the absence of any multicollinearity between them. This allows for the performance of multiple regression analysis.

4.4 Empirical results and analysis

4.4.1 Analysis of the mediating effect of intensity of technological innovation

Table 1 Results of empirical test of the mediating effect of technological innovation intensity

<table>
<thead>
<tr>
<th>explained variable</th>
<th>HLH (Tech)</th>
<th>Tech</th>
<th>HLH (Tech)</th>
<th>GJH (Tech)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>Model 1</td>
<td>Model 2</td>
<td>Model 3</td>
<td>Model 1</td>
</tr>
<tr>
<td>Environ</td>
<td>-0.165***</td>
<td>0.473***</td>
<td>-0.126**</td>
<td>0.225***</td>
</tr>
<tr>
<td>Tech</td>
<td>-0.083</td>
<td></td>
<td>-0.228***</td>
<td>0.411***</td>
</tr>
<tr>
<td>Income</td>
<td>-0.118***</td>
<td>0.226***</td>
<td>-0.099*</td>
<td>0.627***</td>
</tr>
<tr>
<td>R2</td>
<td>0.734</td>
<td>0.743</td>
<td>0.735</td>
<td>0.772</td>
</tr>
<tr>
<td>F</td>
<td>15.566***</td>
<td>145.954***</td>
<td>11.989***</td>
<td>174.862***</td>
</tr>
<tr>
<td>DW</td>
<td>1.952</td>
<td>1.967</td>
<td>1.956</td>
<td>2.028</td>
</tr>
</tbody>
</table>

Note: * denotes p<0.1; ** denotes p<0.05; *** denotes p<0.01.

Data source: Compiled from Eviews 6.0 regression results.

As can be seen from Table 1, the greater the environmental supervision, the more reasonable the industrial structure, and therefore can be tested in the next step. As the intensity of environmental regulation increase, it will increase the technological innovation intensity of enterprises, and therefore proceed to the next test. However, the coefficient of technological innovation intensity is not statistically significant in column (3). According to Wen et al.’s methodology, the Sobel test is necessary. The Sobel test formula indicates that the Z-statistic is significant. Further test
reveals that the significance is reduced. Thus, it can be observed that the intensity of technological innovation plays the role of transmission of the intensity of environmental regulation affecting the rationalization of industrial structure. As the intensity of environmental regulation increases, the industrial structure will also become advanced, and therefore can be tested in the next step. As the intensity of environmental regulation increase, it will increase the technological innovation intensity of enterprises, so the next test can be continued. There is a full mediating effect of technological innovation intensity. Thus, it can be observed that in the process where the intensity of environmental regulation affects the advancement of industrial structure, the intensity of technological innovation plays a mediating role.

5. Further Analysis

5.1 Test for mediating effect by region

The indirect impact of environmental regulatory intensity on industrial structure optimization in the eastern, central, and western areas will be further explored in this study, and the empirical test findings are presented in Table 2.

Table 2 Results of the empirical test on the mediating effect of technological innovation intensity in the eastern, central and western regions

<table>
<thead>
<tr>
<th>explained variable</th>
<th>eastern region</th>
<th>central region</th>
<th>western region</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
<td>Model 3</td>
</tr>
<tr>
<td>Environ</td>
<td>-0.203***</td>
<td>0.764***</td>
<td>0.104</td>
</tr>
<tr>
<td>Tech</td>
<td>-0.402**</td>
<td>-0.719***</td>
<td>-0.402**</td>
</tr>
<tr>
<td>R²</td>
<td>0.748</td>
<td>0.903</td>
<td>0.751</td>
</tr>
<tr>
<td>F</td>
<td>13.439***</td>
<td>188.315***</td>
<td>11.707***</td>
</tr>
<tr>
<td>DW</td>
<td>2.042</td>
<td>1.905</td>
<td>2.171</td>
</tr>
</tbody>
</table>

Note: * denotes p<0.1; ** denotes p<0.05; *** denotes p<0.01.

Data source: Compiled from Eviews 6.0 regression results.
From Table 2, it can be observed that as the intensity of environmental regulation increases, the industrial structure becomes more reasonable, so the next step of testing can be carried out. The more stringent the environmental regulation, the higher the level of technological innovation of the enterprise, so the next test can be continued. The regression coefficient of environmental regulation intensity is not significant, but the regression coefficient of technological innovation intensity is significant. The Sobel test is required. Based on the Sobel test formula, the Z-statistic is significant. It can be seen that the intensity of technological innovation in the eastern region plays the role of transmission of the intensity of environmental regulation affecting the rationalization of industrial structure. The intensity of environmental regulation inhibits the advancement of industrial structure and therefore can be tested in the next step. The technological innovation intensity of enterprises will increase as environmental regulation becomes more stringent, and therefore proceeds to the next test. There is a full mediating impact of technological innovation intensity. Thus, it can be observed that the intensity of technological innovation plays a role of transmission of the intensity of environmental regulation affecting the advancement of industrial structure.

The industrial structure of the central region has become more unreasonable with the increase of environmental regulation, and thus the next test can be conducted. As environmental regulation becomes more stringent, the level of technological innovation within enterprises will increase, so the next test can be continued. At the level of 0.01, there is a significant and positive correlation between the coefficients of technological innovation intensity and environmental regulation intensity, but the coefficient of environmental regulation intensity in column (3) is significantly lower than the coefficient in column (1), and the Sobel Z value is statistically significant, indicating that technological innovation intensity plays a transmission role of environmental regulation intensity affecting industrial structure rationalization. The intensity of environmental regulation inhibits the advancement of industrial structure and therefore can be tested in the next step. The intensity of technological innovation of enterprises will increase with the increase of environmental regulation, and therefore proceeds to the next test. There is a full mediating effect of technological innovation intensity. Thus, it can be observed that the intensity of technological innovation plays a role of transmission of the intensity of environmental regulation affecting the advancement of industrial structure.

The following are the test results for the western region. The amount of technical innovation in the western area is not significantly impacted by environmental regulation intensity. In the western region, there is little impact of environmental regulation intensity on industrial structure optimization.

6. Conclusions

The empirical study's findings indicate that: (1) the intensity of technological innovation in the national, eastern and central regions plays a role in the transmission of the intensity of environmental regulation affecting the rationalization and advancement of industrial structure. However, there is little correlation between the intensity of environmental regulation and the rationalization and development of the industrial structure in the western region.

Acknowledgement. This project is financed by 2022 Social Science Foundation of Shaanxi Province (2022D026).
References

Research on Comprehensive Evaluation of Railway Technical Regulation System Based on FAHP

Qian Li¹, Junting Lin¹*, Zhuo Cai², Yang Wang³

*Corresponding author: linjt@lzjtu.edu.cn

School of Automation & Electrical Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China¹
School of Economics & Management, Lanzhou Jiaotong University, Lanzhou 730070, China²
China Academy of Railway Sciences Corporation Limited (CARS), Beijing 100081, China³

Abstract: To improve the applicability, scientificity, and standardization of the railway technical regulation system, a comprehensive evaluation of the railway technical standard system is required to determine the next optimization direction. In light of the current state of the Chinese railway technical regulation system, this paper provides a comprehensive evaluation of the system using the fuzzy analytic hierarchy process (FAHP). It is suggested that in the future, we can improve the supervision level of the railway technical regulations by beginning with the establishment of deep integration management regulations for process flow and infrastructure.

Keywords: Railway Technical Regulations, Fuzzy Analytical Hierarchy Process (FAHP), Comprehensive Evaluation, Full Lifecycle Management.

1 INTRODUCTION

The railway is a strategic and critical national infrastructure that plays an important role in the national economy and people's livelihood. China's railways have advanced rapidly in recent years, particularly the high-speed lines that have been put into service in large numbers. The railway network's scale and structure are growing larger and more complex. China's railway operating mileage will have exceeded 150,000 km by the end of 2021, with high-speed railway mileage exceeding 40,000 km, ranking first in the world[1]. Simultaneously, numerous new technologies and equipment have been implemented in the railway, and the corresponding rules and regulations are constantly reforming and innovating.

The railway technical regulations define the procedures and methods of railway work, provide the necessary foundation for railway workers' work, and regulate their behavior, which is an important means of realizing safe railway transportation production management. To further improve the applicability, scientificity, and standardization of the railway technical regulations system, it is necessary to establish a sound railway technical regulations evaluation system, conduct a comprehensive evaluation from the standpoint of a clear hierarchy, reasonable structure, standard content, and scientific rigor, and promote the construction of railway technical standards[2].
2 RESEARCH STATUS OF CHINA'S RAILWAY TECHNICAL REGULATIONS

2.1 Railway Technical Regulation System

China's railway technical regulation system has formed a management system with *Railway Technical Management Regulation and Rules for High-speed Railway Operating Organization* as the core rules, and the specific rules of China Railway Group's professional rules, Railway Bureau group company and station section technical rules as the specific rules[3-4]. According to the professional management, according to the nature of the technical rules and regulations, it is divided into system technical rules and regulations, individual technical rules, and according to the professional category of technical rules and regulations, it is divided into train depot, civil engineering, signaling, maintenance, vehicle, communication, vehicles, power supply, freight, passenger transportation, information and other regulations[5]. Figure 1 shows the railway technical regulation system.

![Figure 1. Railway Technical Regulation System.](image)

3 ESTABLISHMENT OF COMPREHENSIVE EVALUATION MODEL FOR RAILWAY TECHNICAL REGULATION SYSTEM

3.1 Fuzzy Analytic Hierarchy Process

The fuzzy Analytic Hierarchy Process (FAHP) is an evaluation algorithm that is based on AHP and fuzzy theory. It is a mathematical evaluation method that takes into account both qualitative and quantitative factors. There are many relevant factors and uncertainties in the evaluation of railway technical regulations that are difficult to divide, making the evaluation
process uncertain and subjective. As a result, the fuzzy analytic hierarchy process can more accurately and reasonably reflect the overall situation of the railway technical regulation system[6].

3.2 Determination of Comprehensive Evaluation Model for Railway Technical Regulation System

In conjunction with railway operation characteristics and risk analysis, and drawing on the indicator setting method of current achievements, three criteria levels are divided: whether the institutional setting and function performance are standardized, the content evaluation of technical regulations, and the effectiveness of technical regulation supervision, and further divided into eight indicator levels. Figure 2 depicts the specific configuration of the evaluation indicator model[7].

4 COMPREHENSIVE EVALUATION METHOD OF RAILWAY TECHNICAL REGULATIONS BASED ON FUZZY COMPREHENSIVE EVALUATION METHOD

4.1 Fuzzy Analytic Hierarchy Process

(1) Construct judgment matrix

First, in conjunction with expert opinions, the 1-9 scale method[8] is used to determine the relative weight value of each factor at each level of the judgment matrix. The judgment matrix is then built, and the consistency test is run to ensure that the analytic hierarchy process (AHP) calculation results are rational. Then there's the n-order judgment matrix $A$ constructed is defined as:

$$A = \begin{bmatrix}
1 & \cdots & a_{in} \\
\vdots & \ddots & \vdots \\
1/a_{in} & \cdots & 1
\end{bmatrix}$$

(1)
(2) Judgment of consistency test

The judgment results are inconsistent because the judgment matrix was constructed subjectively. If the inconsistencies exceed a certain threshold, the judgment results become untrustworthy, so consistency must be checked. The consistency index $CI$ is used to express the judgment matrix’s consistency. The following is the formula:

$$CI = \frac{\lambda_{\text{max}} - n}{n - 1}$$

To compare $CI$ values, the "Random Consistency Indicator" $RI$ is commonly used to measure $CI$ indicators. Table 1 displays the general $RI$ values.

<table>
<thead>
<tr>
<th>n</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>$RI$</td>
<td>0</td>
<td>0</td>
<td>0.58</td>
<td>0.94</td>
<td>1.12</td>
<td>1.24</td>
<td>1.32</td>
<td>1.41</td>
<td>1.45</td>
</tr>
</tbody>
</table>

The results of the comparison are expressed by the unfixied engagement ratio $CR$, as shown in Formula (3):

$$CR = \frac{CI}{RI}$$

The use of $CR$ value to judge the accuracy of the degree of fit is referred to as consistency inspection. When $CR \leq 0.1$, the consistency of the judgment matrix is passed. Otherwise, the value of relative importance in the judgment matrix must be adjusted until the consistency inspection passes.

4.2 Determination of Subordination Degree and Evaluation Level

Following the determination of the indicators, experts will assign the corresponding evaluation grade to each factor in the evaluation model based on the actual situation. Following the scoring of all experts, the frequency of each project grade will be calculated, and the membership degree will be obtained following the normalized calculation, in order to establish a single factor evaluation matrix. $G=\{G1,G2,G3,G4,G5\}=\{I \text{ (very poor), II (poor), III (general), IV (good), V (very good)}\}$ are the five evaluation levels.

4.3 Fuzzy Comprehensive Evaluation

(1) First level fuzzy comprehensive evaluation

Because the factors of the second layer constrain the factors of the first layer, the multifactor evaluation of the second layer will be influenced by the factors of the first layer. As a result,
$R_u$ is chosen as the second layer's independent factor judgment matrix, and $R_u$ is shown in equation (4).

$$
R_u = \begin{bmatrix}
    r_{11} & r_{12} & \cdots & r_{1p} \\
    r_{21} & r_{22} & \cdots & r_{2p} \\
    \vdots & \vdots & \ddots & \vdots \\
    r_{n1} & r_{n2} & \cdots & r_{np}
\end{bmatrix}
$$

(4)

The first level of fuzzy comprehensive evaluation is:

$$
B_{ji} = Z_{ji} \cdot R_u
$$

$$
= (z_{11}, z_{12}, \ldots, z_{np}) \begin{bmatrix}
    r_{11} & r_{12} & \cdots & r_{1p} \\
    r_{21} & r_{22} & \cdots & r_{2p} \\
    \vdots & \vdots & \ddots & \vdots \\
    r_{n1} & r_{n2} & \cdots & r_{np}
\end{bmatrix}
$$

$$
= (b_{11}, b_{12}, \ldots, b_{1p})
$$

(5)

Where, $Z_{ji} = [Z_1, Z_2, \ldots, Z_n]^T$ can represent the weight vector of the $i$-th index in the $j$-th level fuzzy comprehensive evaluation process, and $i$ is the $i$-th evaluation index.

5 MODEL APPLICATION AND EVALUATION RESULT ANALYSIS

This paper invited ten experts in relevant fields and technicians from the Chinese Academy of Railway Sciences to form an expert group to investigate the current state of the railway technical regulation system, grade and evaluate the questionnaire issued, and finally collect the questionnaire for summary, which is used to construct the independent factor judgment matrix of formula (4).

5.1 Construct Judgment Matrix and Perform Consistency Check

The expert group is asked to compare the relative importance of each element in pairs by constructing the judgment matrix shown below using the 1-9 scale method. The weight of each element is then calculated by MATLAB using the above method, and the consistency is checked. Details can be found from Table 2 to Table 5.

| Table 2. Consistency check table of judgment matrix of primary indicators. |
|-----------------------------|----------------|----------------|---|
|                             | $B_1$ | $B_2$ | $B_3$ | $Z_i$ |
| $B_1$                       | 1     | 1/2   | 3     | 0.320 |
| $B_2$                       | 2     | 1     | 4     | 0.558 |
| $B_3$                       | 1/3   | 1/4   | 1     | 0.136 |
The calculated value is 0.0158, less than 0.1, it passes the consistency test.

Table 3. Consistency inspection table of criteria layer (B1).

<table>
<thead>
<tr>
<th>C11</th>
<th>C12</th>
<th>C13</th>
<th>Zi</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>4</td>
<td>0.558</td>
</tr>
<tr>
<td>1/2</td>
<td>1</td>
<td>3</td>
<td>0.320</td>
</tr>
<tr>
<td>1/4</td>
<td>1/3</td>
<td>1</td>
<td>0.136</td>
</tr>
</tbody>
</table>

The calculated value is 0.0158, less than 0.1, it passes the consistency test.

Table 4. Consistency inspection table of criteria layer (B2).

<table>
<thead>
<tr>
<th>C21</th>
<th>C22</th>
<th>C23</th>
<th>Zi</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1/3</td>
<td>2</td>
<td>0.239</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>4</td>
<td>0.625</td>
</tr>
<tr>
<td>1/2</td>
<td>1/4</td>
<td>1</td>
<td>0.136</td>
</tr>
</tbody>
</table>

The calculated value is 0.0157, less than 0.1, it passes the consistency test.

Table 5. Consistency inspection table of criteria layer (B3).

<table>
<thead>
<tr>
<th>C31</th>
<th>C32</th>
<th>Zi</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1/3</td>
<td>0.25</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0.75</td>
</tr>
</tbody>
</table>

Because Table 5 is a second-order matrix, the consistency test is not required, and all of the matrices listed above pass it.

5.2 First Level Fuzzy Comprehensive Evaluation

The basic data scored by experts and the application of Equation 4 and Equation 5 are used for the first level fuzzy comprehensive evaluation.

(1) Whether the institutional setup and function performance are standardized

\[
Z_{11} = (0.558 \ 0.320 \ 0.122)
\]

\[
R_{11} = \begin{bmatrix}
0 & 0.2 & 0.5 & 0.3 & 0 \\
0.3 & 0.2 & 0.4 & 0.1 \\
0.1 & 0.1 & 0.5 & 0.3
\end{bmatrix}
\]

\[
B_{11} = Z_{11} \cdot R_{11} = (0 \ 0.2198 \ 0.3552 \ 0.3564 \ 0.0686)
\]
(2) Content evaluation of technical regulations

\[
Z_{12} = (0.239 \ 0.625 \ 0.136)
\]

\[
R_{12} = \begin{bmatrix}
0 & 0 & 0.3 & 0.5 & 0.2 \\
0 & 0.1 & 0.4 & 0.4 & 0.1 \\
0 & 0.1 & 0.3 & 0.4 & 0.2
\end{bmatrix}
\]

\[
B_{12} = Z_{12} \cdot R_{12} = (0 \ 0.0761 \ 0.3625 \ 0.4239 \ 0.1375)
\]

(3) Effectiveness of technical regulation supervision

\[
Z_{13} = (0.25 \ 0.75)
\]

\[
R_{13} = \begin{bmatrix}
0 & 0.3 & 0.4 & 0.3 & 0 \\
0 & 0.3 & 0.4 & 0.2 & 0.1
\end{bmatrix}
\]

\[
B_{13} = Z_{12} \cdot R_{13} = (0 \ 0.3000 \ 0.4000 \ 0.2250 \ 0.0750)
\]

5.3 Fuzzy Comprehensive Evaluation of Target Level

The fuzzy evaluation matrix of the target level is:

\[
\begin{bmatrix}
B_{11} \\
B_{12} \\
B_{13}
\end{bmatrix} = \begin{bmatrix}
0 & 0.2198 & 0.3552 & 0.3564 & 0.0686 \\
0 & 0.0761 & 0.3625 & 0.4239 & 0.1375 \\
0 & 0.3000 & 0.4000 & 0.2250 & 0.0750
\end{bmatrix}
\]

As shown in Table 2, the weight vector of the target layer is:

\[
Z_2 = (0.320 \ 0.558 \ 0.122)
\]

\[
B_2 = Z_2 \cdot R_2 = (0 \ 0.1494 \ 0.3647 \ 0.3780 \ 0.1078)
\]

To sum up, the results of the fuzzy comprehensive evaluation of the railway technical regulation system are as follows:

\[
G = \begin{bmatrix}
I & II & III & IV & V \\
0 & 0.1494 & 0.3647 & 0.3780 & 0.1078
\end{bmatrix}
\]

According to the principle of maximum subordination degree, the railway technical rules and regulations system's fuzzy comprehensive evaluation level is good, the evaluation system has individual loopholes, and the risk level is low.
6 CONCLUSIONS

This paper conducts a comprehensive evaluation of the railway technical regulation system from the perspective of the railway technical regulation system's construction, as well as the current research situation of the railway technical regulations. The evaluation results show that the railway technical regulation system has a good fuzzy comprehensive evaluation grade, but the effectiveness of technical regulation supervision is relatively weak, and appropriate protective measures are needed to reduce the overall risk level. In the future, we can begin by establishing deep integration management rules of process flow and information flow, and comprehensively strengthen the lossless forward transmission and feedback optimization mechanism of information flow, so that we can realize standardized management of information flow throughout the life cycle based on process flow management, and effectively improve the supervision level of railway technical regulations.

ACKNOWLEDGEMENT This work was supported by the National Natural Science Foundation of China (Grant No. 52162050), and the scientific research fund of China Academy of Railway Sciences (Grant No. 2021YJ205).

REFERENCES

Research on New Media Precision Marketing Method Based on Big Data Information Automatic Push

Yi Zhou¹, a, Yu Chen¹

304810996@qq.com¹

Chongqing College of Architecture and Technology, 401331, Chongqing, China¹

Abstract: Aiming at the problems of low precision and overload of commodity push system, this paper proposes a precise marketing push model based on user portrait feature information. Based on the analysis results and scoring matrix of users' historical behaviors, the user's interest model is constructed, and it is transformed into the user's label model by using label rules. At the same time, the Thrift framework interface is called according to a variety of mixed rules, and the user's preferred push commodity list is returned, so that an accurate marketing push system with excellent performance and low delay time is realized. The simulation results show that the precision marketing push system in this paper uses NDCG algorithm to optimize the parameters, which has high accuracy and stability.

Keywords: Marketing Push, Intelligent Analysis, System Design, Simulation Experiment.

1 INTRODUCTION

It has become an important research topic in the field of intelligent new media information service that how to discover the information that users are interested in from a large number of new media information resources with complex contents and push it to users in time, so that users can get useful information in line with their interests in time. With the help of increasingly mature cloud computing technology, a smart new media information push system based on cloud platform is constructed, which fully integrates massive information resources, realizes the analysis and mining of massive data, and pushes new media information that meets users' personalized needs by combining their information needs, psychological tendencies and behavior habits [¹].

With the rapid development of e-commerce, new network marketing methods have gradually become an important part of e-commerce platform system. In the past 20 years, the number of goods sold by e-commerce companies has been increasing at a high speed every year, which directly leads to the difficulty for users to query the goods they really need. With the large-scale application of mobile Internet technology, this problem has become increasingly prominent. To solve the above problems, the concept of big data with data mining and other technologies as its core has been put forward [²]. However, it is difficult for big data technology to completely solve the above problems at present, that is, the accuracy of data mining technology is low. With the gradual expansion of its application scope, the push systems of many e-commerce
platforms have the problems of low push accuracy and overload of goods. The root is that the design of data mining algorithm of the system is not reasonable enough, and it is difficult to capture users' complex interests, so it is difficult to push the products that users really need [3].

With the continuous expansion of information system business scale and refined management, the increase of information system load is more and more obvious, and the requirements for the reliability and stability of information system are higher and higher. In order to improve the reliability and stability of the information system, Liu and H established a safe and stable performance analysis system. Therefore, the throughput, response time and other performance indicators of the information system are obtained first. Second, establish a performance baseline between load, response time, and resource consumption [4]. In order to improve the marketing effect of e-commerce products, this paper constructs an e-commerce product marketing model based on machine learning algorithm, Cui, F. In addition, this paper studies the classical reinforcement learning algorithm Q learning, and proposes an improved Q learning algorithm [5].

2 DEMAND ANALYSIS

A large number of technology companies such as Microsoft and Amazon have gradually developed personalized push software based on user portraits and interests. Aiming at the problems of low precision and commodity overload of these softwares, this paper analyzes the functional and non-functional requirements of marketing push system in detail.

2.1 Functional Analysis

From the functional point of view, the traditional marketing push system has two serious problems, namely, poor readability of users' data and frequent loading and unloading of goods. To solve these two problems, the push system based on user portrait features should meet the following two main functional requirements:

(1) Good push content, that is, mining the user's preference information from the user's massive behavior data with the greatest strength and degree, and forming push data in a certain format;

(2) As it is difficult for non-technical personnel to extract information from pure data, the push system should realize the visualization of push data, so as to facilitate the operation of data summary and life cycle query by system maintenance personnel.

2.2 Non-Functional Analysis

With the widespread use of mobile Internet terminals, push systems should meet the following non-functional requirements:

(1) The resources occupied by the recommender system are reduced as much as possible, and the recommender system has to deal with massive behavior data, so the CPU and memory of the system are under tremendous pressure. At this time, how to ensure that the system takes up less resources is an urgent problem to be solved.

(2) The recommended system should have a high throughput rate and reduce the error rate of data mining. Because the recommendation system mainly connects the data end and the calling
end of the system by interface, the system will frequently exchange a large amount of data through interface. At this time, within the carrying capacity, how to realize the operation of the system with low error rate and high data throughput rate is a problem that needs to be considered in the system design.

Based on the above demand analysis, this paper puts forward an accurate marketing push system of goods based on the portrait characteristics of users, which consists of subsystems such as user portrait model and personality push platform [6].

2.3 User Portrait Model

In marketing push system, user portrait is a computing model that depicts the whole situation and realizes labeling. In order to realize the personalized push of goods, this paper designs several modules of the user portrait model, and the specific architecture is shown in Table 1.

<table>
<thead>
<tr>
<th>TABLE I. USER PORTRAIT MODEL MODULE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>User portrait model</strong></td>
</tr>
<tr>
<td><strong>Label display</strong></td>
</tr>
<tr>
<td>Visualization display</td>
</tr>
<tr>
<td>Preference data display</td>
</tr>
<tr>
<td>Market data display</td>
</tr>
<tr>
<td>Life cycle display</td>
</tr>
<tr>
<td>Build label model</td>
</tr>
<tr>
<td>Portrait feature description</td>
</tr>
<tr>
<td>Constructing interest model</td>
</tr>
<tr>
<td>User data acquisition</td>
</tr>
</tbody>
</table>

As can be seen from Table 1, the user portrait model consists of portrait feature description and visual display. Among them, the former is the core of the portrait model, while the latter is mainly responsible for the display of the model. This paper mainly elaborates the methods of portrait feature description, and the specific contents are as follows:

User data acquisition. In the user portrait model, collecting user data is the most basic and important step in building the model. JavaScript is often used to collect web page data and store it in HDFS file system. The specific architecture is shown in Figure 1.
(2) Build an interest model. Building an interest model is also a key step in building a user portrait model. This step uses the user's basic and behavioral data to guess their personal preferences, and shows the user's interest through the corresponding scoring algorithm.

(3) Build a label model. In the user portrait model, the tag model needs to analyze the necessary user behaviors, preferences and basic attributes according to the user interest model. In these analysis processes, the model needs to analyze a large number of user data indicators [7].

### 2.4 Personality Push Platform

After the user portrait model is built, the system also needs to implement a push platform with strong personality. In the implementation of this platform, this paper divides it into Task layer and Service layer: Task layer is responsible for extracting user interest model data, and Service layer is responsible for implementing the encapsulated data interface. It is elaborated as follows:

**Task layer.** In the implementation of Task layer, when the platform obtains the user's interest model, the push platform needs to complete three specific processing operations. ① The push platform removes the expired off-shelf products by calling the interface of the current product; ② The push platform needs to obtain the user's historical push data, so as to filter the pushed products.

**Service layer.** In the implementation of the Service layer, the push platform needs to provide the corresponding mixed operation rules for the interest model in the form of interface, and its specific operations include: ① The platform needs to set the operation positions of commodities, add the operation positions of other competing products with higher commercial value, and realize the configurability of commodity operation positions; The push platform must be able to tolerate a certain error rate, which is because it is difficult to keep the commodity information in real life in an accurate state; ③ The push platform mainly pushes specific commodity information to users through Thrift interface [8].

### 3 SIMULATION EXPERIMENT

Based on the above demand analysis, this paper proposes a precision marketing push system for goods based on the user's portrait characteristics, which is composed of user portrait model, personalized push platform and other subsystems. Its application architecture is shown in Figure 2.

It can be seen from Figure 2 that, first of all, in the user portrait model, the system uses a large number of user basic data and behavior data to extract and form the unique tags and attributes of users, so that all user groups can be divided in detail; Secondly, the personalized push platform extracts the tag model in the user profile from Hive warehouse tool, and provides it to the calling device in the form of interface call after filtering.
In order to verify the accuracy and stability of marketing push system, this paper uses Web, data collection and Hadoop distributed servers to test the accuracy of commodity push. In the simulation process, the hardware and software environment of the push system are set up. Among them, Huawei cloud server is used in the hardware, and its memory is 8 GB and hard disk is 80 GB. This paper adopts Linux distribution, Redis 2.6 and Hive database, and Hadoop 1.2 distributed server software. In the specific simulation process, this paper randomly selects 10,000 user data from the published real business test data set [9]. Through five consecutive tests and calculation of user satisfaction rate, the traditional push system is simulated and compared with the push system proposed in this paper, and the corresponding simulation results are obtained, as shown in Table 2.

<table>
<thead>
<tr>
<th>Test times</th>
<th>Traditional system</th>
<th>This paper system</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.541</td>
<td>0.978</td>
</tr>
<tr>
<td>2</td>
<td>0.348</td>
<td>0.986</td>
</tr>
<tr>
<td>3</td>
<td>0.426</td>
<td>0.968</td>
</tr>
<tr>
<td>4</td>
<td>0.208</td>
<td>0.976</td>
</tr>
<tr>
<td>5</td>
<td>0.436</td>
<td>0.972</td>
</tr>
</tbody>
</table>

As can be seen from Table 2, the user satisfaction rate of the traditional push system is between 0.208 and 0.541 during many tests; The user satisfaction rate of the push system proposed in this paper is between 0.968 and 0.986. It shows that the user satisfaction rate of the proposed marketing push system is significantly better than that of the traditional marketing push system [10-15].

4 CONCLUSIONS

Based on the characteristics of user portraits, this paper proposes a high-precision marketing push system. For a large number of e-commerce platform users, this system can push some kind of commodity information that users really prefer, thus saving advertising costs and enhancing the commercial value of e-commerce platform. However, there are still some
shortcomings in the system proposed in this paper. For example, the push weight calculation of personalized recommendation platform is not accurate enough, which needs further optimization and calibration. This problem will be solved in the follow-up research.

REFERENCES

A Design Method of Logistics Support Simulation Support Software

Yong Sun¹,a, Shaopan Zhang¹, Guangzhao Lu¹, Jinqing Zhao¹, Jianhui Tian¹, Jirong Xue¹

sunyong1984@gmail.com⁶

China Research and Development Academy of Machinery Equipment Beijing, China¹

Abstract: In the process of system of systems simulation, logistics equipment support simulation is an important part. To meet the needs of logistics support simulation, this paper proposes a design method of logistics support simulation support software, and implements corresponding software based on this method. This method includes logistics support simulation initialization and data receiving module, logistics support equipment simulation module, logistics support command simulation module, health support simulation module, transportation simulation module, fuel and material support simulation module. The logistics support simulation data output and record module is composed. Based on this method, the logistics support simulation problem can be solved. Combined with the actual project application, good results have been achieved.

Keywords: Logistics Support, Simulation Support, Health and Epidemic Prevention, Field First Aid.

1 INTRODUCTION

Logistics equipment support simulation is a common field in the system of systems simulation. At the same time, it is also a difficult problem because it involves a wide range of disciplines, multiple equipment and complex process. Logistics support, as the basic work of system of systems simulation, has been widely studied at home and abroad. The existing research is roughly divided into two categories, one is to study the performance of support equipment [¹, ²], the other is to combine advanced computer technology and artificial intelligence technology [³, ⁴], but they have not achieved particularly good results. Aiming at the actual needs of system simulation, this paper proposes a design method of logistics support simulation support software for logistics support equipment simulation, logistics support command simulation, medical support simulation, oil and material support simulation and other problems, and implements the software based on this method, and carries out relevant applications in combination with actual projects. The following are introductions.
2 LOGISTICS SUPPORT SIMULATION INITIALIZATION AND DATA RECEIVING MODULE

1) If the computer startup information released by the evaluation control module is received, start the computer, and release the response information after startup.

2) If the data distribution information released by the evaluation control module is received, the received data is stored in the computer hard disk and the response information is released.

3) If you receive the startup program and data initialization information released by the evaluation and control module, start the simulation system and read the relevant data, and release the response information after the above work is completed. Call the plane normal vector simulation basic model, calculate the normal vector of each surface defined in each equipment entity, call the fault simulation basic model, simulate the fault time of the combat unit, store the time in a fault data file, and add a data item (the time required to check the fault of the combat unit, take 3 minutes).

4) If you receive the Federation information released by the evaluation and control module, complete the Federation operation.

5) If the clock synchronization information released by the evaluation control module is received, adjust the time of the computer to be the same as that of the evaluation control computer. If the time service information released by the evaluation control module is received, store the operation time at the beginning of the simulation in the computer hard disk and release the response information.

6) If the simulation start information is received, release the response information to the simulation management computer and run the software.

7) If the simulation end information released by the evaluation control module is received, the operation of the simulation system will be terminated.

8) If the simulation pause, continue and end information released by the evaluation control module is received, RTI will control the pause, continue and end of the simulation system.

9) If the federate joining and exiting information released by the evaluation and control module is received, the federate joins or exits the simulation.

10) If the meteorological environment information released by the evaluation and control module is received, the information is stored in the array.

11) If the computer shutdown information is received, issue the response information and shut down the computer.

12) If the equipment entity status information is received, replace the original information with the new information.

13) If you receive laser ranging information for yourself, judge whether the combat unit has laser warning and smoke screen function (query the comprehensive protection performance data of the combat unit according to the combat unit model). If so, call the basic smoke simulation model. If the smoke is successful, release the entity status information.
14) If the direct hit target information is received, judge whether the target is the combat unit. If so, query whether the combat unit has active protection capability (soft kill active protection capability or hard kill active protection capability) from the comprehensive protection performance data file of the combat unit. If not, call the basic simulation model of damaging armored target during direct hit. If the combat unit has active protection capability, call the basic simulation model of active protection. If the active protection fails, call the basic simulation model of armor target damage in direct hit.

15) If the information of missed target is received, judge whether the target is the combat unit. If yes, call the basic simulation model of armor target damage in case of missed target.

16) If the combat unit is damaged, issue entity status information and damage information. If there are no casualties or slightly injured personnel, it is considered that the combat situation of the combat unit can be reported, and the damage information is stored in a damage data file.

17) If the camouflage information is received, judge whether the support object is the combat unit. If so, release the entity status information.

18) Issue entity status information if it receives the decontamination information of this combat unit.

19) If receiving the equipment repair information and equipment rescue information of this combat unit, issue the entity status information.

3 LOGISTICS SUPPORT EQUIPMENT SIMULATION MODULE

This module is a time cycle of federal members. Set the current simulation time as t, and complete the following operations for K combat units in the combat unit deployment data file within each simulation cycle \( \triangle T \):

Query the fault data file of the combat unit, if any, when the simulation time t meets:

\[ t < \text{failure time of the combat unit} < t + \triangle T \]

Publish entity status information when.

Query the fault data file of the combat unit, if any, when the simulation time t meets:

\[ t < \text{the failure time of the combat unit} + \text{the time required to check the failure of the combat unit} < t + \triangle T \]

According to the fault information, fill in the battle information and publish it to the communication network evaluation module.

Query the damage data file of the combat unit, if any, when the simulation time t meets:

\[ t < \text{the damage time of the combat unit} + \text{the time required to check the damage of the combat unit} < t + \triangle T \]
According to the damage information, fill in the war situation information and publish it to the communication network evaluation module.

If the combat unit has positioning function (query the performance data file according to the combat unit model) and the combat unit is not damaged or faulty, set the distribution cycle of positioning information as $\Delta T_1$ (query the combat unit network deployment data file). $\Delta T_1$ takes the integer multiple of $\Delta T$, and make $n = \frac{\Delta T_1}{\Delta T}$. Within $n$ simulation cycles, call the basic model of equipment positioning simulation to calculate the combat unit position output by the positioning system, fill in the data structure of positioning information and publish it to the communication network evaluation module \(^5\).

Call the basic model of maneuver simulation, calculate the grid number $I$ and $j$ occupied by the time position of combat unit $T$, compare it with $I$ and $j$ in the last entity state, and release the entity state information if there is any change.

Use the deployment data and maneuver data of the combat unit to calculate the entity state information at time $t$, and update the entity state information together with the information of the previous entity state modification.

If the type is logistics support command equipment, call the logistics support command simulation module.

If the entity type is medical support equipment, call the medical support simulation module.

If the entity type is transportation support equipment, call the transportation simulation module.

If the entity type is oil and material support equipment, call the oil and material support simulation module.

Query the intelligence and command and control information data file. If the simulated equipment releases unformatted intelligence and command and control information in this simulation cycle, release the information to the communication network evaluation module.

4 LOGISTICS SUPPORT COMMAND SIMULATION MODULE

The logistics support command simulation module is composed of logistics support command and control model.

The logistics support command and control model mainly completes three functions: command receiving, command analysis and command distribution. The logistics support commands include tactical maneuver commands, field operation commands, health and epidemic prevention commands, rescue commands, casualty transportation commands, material supply commands, transportation plan commands and oil supply commands \(^6\).

It can realize the command function simulation of the commander of the combat unit, including receiving the combat command and support information issued by the superior command entity, decomposing and processing them, forming the determination of the current level, issuing them to the corresponding equipment, and reporting the combat situation to the superior.
Command receiving, read the commands distributed to yourself from the global linked list of the engine, and save the commands locally.

Command analysis: analyze the received command to obtain the sender, sending time, task type, task start time, task parameters and other information.

Command distribution: send commands to each task execution mode according to the parsed information.

The logistics support command and control model is composed of several functional modules, such as command receiving, command parsing and command distribution.

Description of typical logistics support command and control model simulation process:
Step 1: initialize data.
Step 2: read the damage information from the damage model. If it has been destroyed, end.
Step 3: receive the support information and analyze the support information.
Step 4: generate a support command, and then issue it to the corresponding support equipment.

5 HEALTH SUPPORT SIMULATION MODULE

The health service support simulation module can realize the simulation of various health service equipment functions, which is composed of health and epidemic prevention vehicle simulation model, field emergency vehicle simulation model, field surgical vehicle simulation model and casualty transport vehicle simulation model [7].

1) Simulation model of sanitation and epidemic prevention vehicle

The simulation model of health and epidemic prevention vehicle mainly completes three functions: data receiving, data analysis and processing and health and epidemic prevention.

Data receiving, receiving the command data distributed by the command and control module.
Command analysis and processing, analyze the received command data, so as to obtain the support location, support unit and other information.
Sanitation and epidemic prevention: carry out sanitation and epidemic prevention in designated areas.

The simulation model of health and epidemic prevention vehicle is mainly composed of three functional modules: data receiving, data analysis and processing and health and epidemic prevention.

2) Simulation model of field ambulance

The simulation model of field ambulance mainly completes three functions: data receiving, data analysis and processing and field first aid.

Data receiving, receiving the command data distributed by the command and control module.
Command analysis and processing, analyze the received command data, so as to obtain the support location, support unit and other information.

Field first aid: conduct field first aid at the designated place.

The simulation model of field ambulance is mainly composed of three functional modules: data receiving, data analysis and processing and field ambulance.

3) Simulation model of field operation vehicle

The simulation model of field operation vehicle mainly completes three functions: data receiving, data analysis and processing and field operation.

Data receiving: receive the command data distributed by the command and control module.

The command analysis process parses the received command data to obtain the support location, support unit and other information.

Field operation, field operation.

The simulation model of field operation vehicle is mainly composed of three functional modules: data receiving, data analysis and processing and field operation.

4) Simulation model of casualty transport vehicle

The simulation model of the wounded transport vehicle mainly completes three functions: data receiving, data analysis and processing and transporting the wounded.

Data receiving, receiving the command data distributed by the command and control module.

Command analysis and processing, analyze the received command data, so as to obtain the support location, support unit and other information.

Transport the wounded and carry out the operation of transporting the wounded.

The simulation model of casualty transport vehicle is mainly composed of three functional modules: data receiving, data analysis and processing and casualty transport.

6  TRANSPORTATION SIMULATION MODULE

The transportation simulation module can realize the simulation of various transportation equipment functions, which is composed of transportation vehicle simulation model.

The transport vehicle simulation model mainly completes three functions: data receiving, data analysis and processing and transportation [8].

Data receiving, receiving the command data distributed by the command and control module.

Command analysis and processing, analyze the received command data, so as to obtain the support location, support unit and other information.

Transport the wounded to the designated area.
The simulation model of transport vehicle is mainly composed of three functional modules: data receiving, data analysis and processing and transportation.

7 OIL AND MATERIAL SUPPORT SIMULATION MODULE

The oil and material simulation module can realize the simulation of the functions of various oil and material supply equipment, which is composed of the oil supply vehicle simulation model.

The simulation model of oil supply vehicle mainly completes three functions: data receiving, data analysis and processing and oil supply.

Data receiving, receiving the command data distributed by the command and control module.

Command analysis and processing, analyze the received command data, so as to obtain the support location, support unit and other information.

Oil supply: supply oil to the designated area.

The simulation model of fuel supply vehicle is mainly composed of three functional modules: data receiving, data analysis and processing and fuel supply.

8 LOGISTICS SUPPORT SIMULATION DATA OUTPUT AND RECORDING MODULE

The logistics support simulation data output and recording module completes the recording of release and order data.

Combined with the actual project, several simulations were carried out. The simulation time and task completion data of each simulation module are shown in the following table.

<table>
<thead>
<tr>
<th>Module name</th>
<th>Simulation time</th>
<th>Task completion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistics support equipment simulation module</td>
<td>43</td>
<td>82%</td>
</tr>
<tr>
<td>Logistics support command simulation module</td>
<td>29</td>
<td>74%</td>
</tr>
<tr>
<td>Health support simulation module</td>
<td>13</td>
<td>90%</td>
</tr>
<tr>
<td>Transportation simulation module</td>
<td>18</td>
<td>77%</td>
</tr>
<tr>
<td>Oil and material support simulation module</td>
<td>26</td>
<td>85%</td>
</tr>
</tbody>
</table>
9 CONCLUSION

According to the needs of logistics support simulation field, this paper proposes a design method of logistics support simulation support software, and realizes the corresponding software based on this method. The method includes logistics support simulation initialization and data receiving module, logistics support equipment simulation module, logistics support command simulation module, health support simulation module, transportation simulation module, oil and material support simulation module. Logistics support simulation is composed of data output and recording module. Based on this method, the problem of logistics support simulation can be solved.

REFERENCES

Design of User Portrait Analysis System Based on E-Commerce Big Data

Yu Chen¹*, Yi Zhou¹
*550417230@qq.com
Chongqing College Of Architecture and Technology, 401331, Chongqing, China¹

Abstract: This paper analyzes the relationship between user portrait and e-commerce, and probes into the construction strategy of user portrait system. From the aspects of precision marketing, user statistics, data mining, building intelligent recommendation system, marketing effect evaluation, improving the operation mode of products and services, user satisfaction management, private customization of products and services, etc., with the continuous growth of user data, the scope and dimensions of user portrait model will be more diverse. According to the data of a shopping mall, multi-dimensional labels such as rule matching, statistics and data mining are implemented to form user portraits. Suggestions are put forward for the development of e-commerce user portrait system.

Keywords: E-Commerce, User Portrait, Analytic System.

1 INTRODUCTION

In the era of developed Internet, e-commerce industry has mushroomed, and now most people will choose to buy what they want online, which will be much more convenient. The rapid popularization of the Internet leads to the generation of a large amount of data, which leads to the problem of network information overload. To solve this problem, the creation of user portrait is the most important step to achieve accurate information recommendation [1]. Among them, the generation of the label of the user portrait is the digital description of the user's characteristics, and the final portrait model is formed mainly by analyzing and mining the behavioral data of the user. When users fill in data or browse the page, they will leave a lot of buried data, which invisibly reflects the diversity of user information. By analyzing the diversified data of users, we can get the general preferences and habits of users, and then label users. Users can generate the best recommendation through big data mining algorithm, which can provide huge data support for some advertising, accurate sending of short messages, network management and other aspects [2]. The development of big data has broken the traditional phenomenon of spreading the net indiscriminately, which not only saves the cost for enterprises, but also brings greater benefits to enterprises while meeting the needs of customers.
2 USER PORTRAIT AND E-COMMERCE

User portrait technology is to collect and label user information with the help of big data. Through big data, abstract important information such as users' social attributes, living habits and consumption habits, and establish a labeling model. From the perspective of e-commerce, based on the information that users fill in on e-commerce websites and the information of users' consumption behaviors, users are depicted with some tags, and the depicted user tags are user portraits. E-commerce is a kind of commodity trading mode that develops commodity trading activities with the help of network platform. E-commerce is the electronization and networking of traditional business activities. E-commerce has been developing with the popularization of the Internet and the increase of the number of Internet users [3]. At present, all industries have e-commerce platforms, which have become a new way for enterprises to organize commodity activities and played an important role in the economic growth of enterprises. The development of e-commerce can not only promote enterprises to generate income, but also change people's daily life style and consumption style. E-commerce has become a new driving force for China's economic development [4].

3 DESIGN OF USER PORTRAIT SYSTEM

3.1 Overview of User Portrait

User portrait is a bridge to depict customers and link users' needs with design. User portrait has a variety of applications in various industries in today's society. In the actual work process, we often use simple words that are close to life to connect the attributes and behaviors of users with the data in the database, so as to establish a relationship. User portrait is actually the virtual image of every user [5]. Its formation depends on and is higher than that of users, and it will not be independent of products and markets. Instead, we can learn about the target user groups of our enterprise through user portraits, and make appropriate adjustments and activity plans for them. This user model can bring intuitive and specific data information to the enterprise for reference. So as to understand users, discover their potential needs, finely locate different users according to different products and even dig out potential users. Especially, the application in the recommendation system is particularly important. Thousands of people refer to recommending different products according to different users. They can also combine the data obtained from advertising companies, manufacturers, express delivery companies, self-media websites, short videos, etc. to dig deep into the characteristics of users and discover their differences and common characteristics, so as to find marketing opportunities and plan directions for users and comprehensively improve the core influence of the company [6].

3.2 User Portrait Application

User portraits can not only help enterprises to understand users more comprehensively, but also subdivide different users and groups. According to different groups, they can make information predictions to analyze users' hobbies and product rising space, and dig out users' needs and pain points in different usage scenarios. Here's an introduction to the application of user portraits in different directions:
(1) Precision marketing

According to the user's portrait, the different needs and interests of the user are analyzed, and then the products needed by the user are pushed pertinently. Introduce the right products to the right users at the right time to realize accurate marketing. For example, the classification of goods is analyzed by the user's shopping records, and similar goods are pushed when the user enters the shopping page next time, which can greatly shorten the user's search time, improve the click-through rate and repurchase rate of the user, and increase the user's dependence and loyalty to websites or software [7].

(2) Industry analysis

After having a large amount of user data, you can know the current development status of the whole industry, the advantages and disadvantages of the products that users like, and the direction provided for the adjustment of enterprises. At the same time, you have a clear understanding of the positioning of your own products in order to improve the products and services again.

(3) User positioning

By observing users' repeated purchases of certain commodities, the characteristics of such user groups and the competitiveness of products can be analyzed, so that similar commodities can be better recommended to such loyal customers, and more clicks and orders can be made.

(4) Advertising business

After knowing the user's preferences through the user's portrait, it can provide detailed data reference for the advertising of merchants. So as to maximize that benefit of the next advertisement [8].

3.3 User Portrait Construction Process

The construction process of user portrait can be divided into the following four steps, as shown in Figure 1:

![Figure 1 User portrait construction process](image)
(1) Source data acquisition. Static data often comes from the information filled in by users themselves in the website. When there is no data, it can be predicted by establishing a model. For example, when the gender information of users is unknown, a model can be established to judge by users' purchasing behavior. Dynamic data refers to the data generated by users' behaviors, such as browsing pages, purchasing goods, collecting, product evaluation, etc. These behavioral data can calculate the customer's brand preference, spending power, order quantity, category purchase ranking and other information.

(2) Data preprocessing. When data is collected, it is generally necessary to do some pre-processing, such as data cleaning and data structuring. While removing some dirty data, we can standardize the structure of the data to facilitate the subsequent better processing and analysis.

(3) Behavior modeling. Through statistics of different dimensions and calculation of algorithm models, users' static data and dynamic data are labeled to express users' interests, needs and preferences.

(4) Creation of user portrait. Deepen the process of behavior modeling, and use prediction, clustering algorithm, machine learning, text mining and other related technologies to complete the mapping of different attributes of users to tags. The completion of user tagging means that the portrait of the user is basically completed.

3.4 Gender Model Construction

In e-commerce websites, it is extremely important to know the gender information of users, because there are great differences between men and women in their shopping needs, interests and preferences, and shopping frequency. Therefore, in the construction of user portraits, gender tag is one of the tags that cannot be ignored. At present, the website can't accurately distinguish the gender of some users because users mistakenly fill in their accounts, don't fill in their gender, or often use their family accounts. Based on this situation, the best way to deal with the website at present is to predict and judge the gender of users according to their behaviors on the website. This paper also builds a gender model based on users' browsing and purchasing behavior. Because of the different needs of men and women, boys often like to search electronic products, men's wear, men's shoes, razors, belts and other commodities on e-commerce websites; Girls usually search for products such as skin care products, high heels and women's clothing [9]. According to the user's search records and click records, this paper uses the improved naive Bayesian classification algorithm based on the improved EM algorithm to predict the user's gender. The gender classification dimensions are divided into male, female and unknown. In order to facilitate the further study of user portraits, this paper adds several index dimensions in Table 1, so as to better understand users:

<table>
<thead>
<tr>
<th>Gender</th>
<th>Index dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Male</td>
<td>Purchase order quantity of male characteristic category</td>
</tr>
<tr>
<td>2. Female</td>
<td>Purchase order quantity of female characteristic category</td>
</tr>
<tr>
<td>3. Unknown</td>
<td>Number of browsing male characteristic categories</td>
</tr>
<tr>
<td></td>
<td>Number of browsing female characteristic categories</td>
</tr>
<tr>
<td></td>
<td>Time spent browsing male characteristic categories</td>
</tr>
<tr>
<td></td>
<td>Time spent browsing female characteristic categories</td>
</tr>
</tbody>
</table>
3.5 City-Level Model

After the user's gender model is constructed, the user's urban model is constructed. In the e-commerce website, the city where the users are located is also an important dimension reference, and the shopping needs of users in different cities are also different. For example, the demand for down jackets in southern cities is not as big as that in northern cities, and the frequency of buying seafood in cities near the sea is not as high as that in other cities. At present, the IP of the network corresponds to the city, so the user's city is usually obtained by the IP address and the receiving address. For users who lack the information of the user's city, the naive Bayes algorithm can also be used to predict the city, and the same algorithm can be used to predict it. Also, this article adds several additional indicator dimensions in Table 2, so as to better understand users:

<table>
<thead>
<tr>
<th>City level</th>
<th>Index dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Line city users</td>
<td>Delivery address</td>
</tr>
<tr>
<td>2 Line city users</td>
<td>IP login address</td>
</tr>
<tr>
<td>3-4 Line city users</td>
<td>Use coupon amount</td>
</tr>
<tr>
<td>5-6 Line city users</td>
<td>Number of coupons used</td>
</tr>
<tr>
<td></td>
<td>Total order amount</td>
</tr>
<tr>
<td></td>
<td>Total number of orders placed</td>
</tr>
</tbody>
</table>

4 APPLICATION OF E-COMMERCE USER PORTRAIT SYSTEM

The application of e-commerce user portrait system has the following aspects. ① Precision marketing. Through the analysis of potential users, on the basis of information technology, we will carry out precise marketing for specific user groups. ② User statistics. Through e-commerce user portraits, users can be counted, and more scientific development goals can be formulated based on relevant data. ③ Data mining. Through e-commerce user portraits, we can deeply mine data and grasp user characteristics, so as to build an intelligent recommendation system. ④ Evaluation of marketing effect. Only by perfecting the operation mode of products and services can the quality of products and services be improved. With the help of user portraits, users can be accurately analyzed, service groups can be effectively positioned, and service levels can be significantly improved. ⑤ User satisfaction management. Through user portraits, e-commerce can effectively manage user satisfaction and formulate improvement measures according to user satisfaction. ⑥ Private customization of products and services. Based on user portraits, e-commerce can effectively grasp the consumption situation of certain consumer groups, so as to realize the personal tailor of products and services. User portrait technology can be applied to the whole process of users' purchase, helping e-commerce to provide better services for users, and helping to enhance the competitiveness of e-commerce [10].
5 Conclusion

In the era of big data, to develop e-commerce, user portrait technology is essential. This technology can accurately grasp the needs of users and provide personalized products and services for users. In the new period, to strengthen the design of user portrait system, we can improve the user portrait technology, strengthen the application of user portrait system, and promote the rapid economic development through technological innovation.

REFERENCES


Abstract: This paper attempts to study the relationship between price-earnings ratio and stock returns. We find that stocks with lower price-earnings ratio can obtain higher returns than stocks with higher price-earnings ratio. This finding applies not only to simple returns, but also to risk-adjusted returns based on CAPM, as well as Fama-French (1993) three-factor model. In addition, to test the ability of PE risk factors in explaining asset prices, this price-earnings ratio factor is added into the Fama-French three-factor model as a new pricing factor. The results show that PE risk factor significantly affect stock returns. This contributes to the asset pricing literature.

Keywords: P/E ratio, Asset Pricing, Fama-French Three Factor Model, CAPM.

1 INTRODUCTION

In the era of digital economy, data has become a new key factor of production. Investors try to find factors related to stock returns with a large number of corporate information datasets. These pricing factors are then used to predict future stock prices and generate investors additional returns. The existing capital asset pricing model (CAPM) propose that stock market premium is the key systematic risk factor that influences individual asset returns. Later, the Fama-French (1993) three-factor model (FF3) propose additional two factors which can explain the stock returns, namely, the size factor and value factor. Relevant literatures show that the price-earnings ratio is very useful for stock valuation [1-3]. Stocks with low P/E ratios usually outperform stocks with high P/E ratios, which means investors can get higher returns by choosing stocks with low P/E ratios.

The purpose of this paper is to study whether the price-earnings ratio can explain the return of the stock, and whether the stock with lower price-earnings ratio can obtain higher return than the stock with higher price-earnings ratio. In addition, we use the long-short investment strategy that we hold stocks with low price-earnings ratio and short-sell stocks with high price-earnings ratio. The evidence shows that such investment strategy could generate significant positive returns. Afterwards, we add the P/E ratio to the FF3 as a new pricing factor, namely, the PE risk factor. We then examine whether the PE risk factor is useful in explaining the return on the portfolio, in addition to the FF3.

The paper is organized in the following way. Section 1 is the introduction. Second 2 reviews relevant literate and discusses the contributions of this paper. The third section describes the sample details and data sources. The empirical methodology is presented in the fourth section. Section 5 analyses the empirical results and discusses the application. Section 6 analyses the new pricing factors. Section 7 concludes.
2 LITERATURE REVIEW

According to Sharpe (1989), Harry Markowitz proposed the concept of (mean-variance) efficient portfolio in 1952. Mean-variance analysis can determine the maximum return at a given risk level or the minimum risk at a given return level. Markowitz's portfolio theory is based on the mean-variance optimization process of finding an effective portfolio. Markowitz (1952, 1959) develop a portfolio model to measure the relationship between risk and expected return. The author shows that under a set of reasonable assumptions, the variance of return is a meaningful measure of portfolio risk [8]. However, when using Markowitz's portfolio theory to evaluate the risk of portfolio, extensive computing power are needed to construct the variance-covariance matrix across various number of assets. The amount of computation required by Markowitz method is one of the reasons that stimulate other investment management methods [7].

Sharpe (1964) and Lintner (1965) propose the CAPM on the model of portfolio choice developed by Markowitz (Fama and French 2004). Furthermore, CAPM adds two factors to the Markowitz model, risk-free rate (RFR) and market risk premium (E(RM) - RFR). According to Reilly and Brown (2011), CAPM redefines the related measure of risk from the total risk defined by capital market line (CML) to only the non-diversifiable part of the total risk. This is termed as the systematic risk, which is called beta. Then, CAPM redefines the expected risk premium per unit of risk by using beta as the relevant systematic risk measure.

However, Fama and French (1996) claim that CAPM did not capture certain risks and therefore could not explain the abnormal returns of some stocks. Therefore, Fama and French proposed a three-factor model to test the average stock returns. The three-factor model adds a size factor and a value factor to the CAPM model. According to Reilly and Brown (2011), SMB aims to capture the risk elements associated with company size, while HML aims to differentiate the risk differences associated with "growth" (i.e., low book-to-market ratios) and "value" (i.e., high book-to-market ratios) companies. However, both price-earnings ratios and book-to-market ratios are commonly used to classify growth and value stocks. Therefore, as it is evidence, P/E ratio plays a crucial part in asset pricing.

Basu (1975) uses the NYSE industrial firms from the COMPUSTAT dataset, the related Investment Return file from the CRSP and a file containing selected financial statement and investment return data for securities subsequently delisted from the NYSE. In addition, this study compute earnings yields of each stock and then rank from minimum to maximum and 5 portfolios are formed. Finally, monthly and cumulative abnormal returns are calculated for each portfolio. This article of Basu (1975) attempts to ascertain whether securities with different P/E ratios are appropriately priced, or whether certain groups can gain abnormal returns. Finally, the results of Basu (1975) shows that low PE ratio portfolios outperform high PE ratio portfolios generating positive and higher risk-adjusted rates of return.

As an extension, Basu (1977) further investigates to determine whether the investment performance of securities is related to their P/E ratios. Basu (1977) uses that same database as Basu (1975) and calculate the P/E ratio of each stock and rank the P/E ratios to form five portfolios. Calculating the monthly rate of return for each portfolio over the past 12 months and then selects a portfolio to buy and hold for the next 12 months. Reordered each year and then reinvested same portfolio with the previous period. The results indicate that the return of a portfolio with low P/E ratio is higher than that of a portfolio with high P/E ratio. Basu (1977)
shows that the behavior of securities prices during the sample period may not be fully described by the efficient market hypothesis, and the portfolio with low P/E ratio does earn higher returns. In Basu (1983), it studies the relationship between P/E ratio and firm size on stock returns. The database includes accounting earnings per share from Compustat and the stock price, return and ordinary shares data from the CRSP monthly return data and all sample stocks must be listed on the NYSE. Basu (1983) calculates the P/E ratio each year and ranks them in ascending order to form five portfolios. In addition, five size portfolios are formed based on market value. By controlling the firm size and P/E ratio respectively, the author constructs the portfolio of yield and market value. Calculate and compare the relationship between risk and return of these portfolios, and finally make a statistical test on their risk-adjusted returns to determine whether there is a significant yield/size effect. The research results confirm that the stocks with high P/E ratio earn higher risk-adjusted returns on average than the stocks with low price-earnings ratio, and even if the experimental control is implemented on the difference in company size, this effect is also significant.

Afterwards, Noda, Martelanc and Kayo (2015) examine whether the return of the portfolio based on the E/P ratio of stocks is significantly different from that predicted based on CAPM model. The sample includes all companies listed on the Sao Paulo Stock Exchange (BOVESPA) from January 1995 to March 2013. The authors sort the stocks according to the market value, B/M and E/P ratio, classifies them into the portfolio, and then calculates the portfolio return. In order to check whether the portfolio composed of E/P ratio has a significantly different return from that provided by CAPM, the ordinary least square regression method is used. The regression intercept, alpha, therefore measures the abnormal return of the stock. In addition, Noda, Martelanc and Kayo (2015) add an earnings/price factor, i.e., high earnings minus low earnings, as a new risk factor in addition to the FF3. The empirical evidence suggests that portfolio with high E/P ratio is often higher than the return predicted by CAPM. E/P risk coefficient is very important to explain the return of portfolio, whether in CAPM or FF3.

3 DATA & VARIABLE

There are three main sources of data for this survey. Monthly securities price data are obtained from the CRSP. We include the stocks of non-financial companies listed on the NYSE, the Amex and the NASDAQ stock exchanges. That is, we exclude companies with four-digit SIC codes between 6000 and 6999, which stands for financial firms. The sample period ranges from 1963 to 2022. Ultimately, we have 3207 firms and 124950 firm-month observations. In addition, company annual financial statements and accounting information are downloaded from the Compustat data center. Furthermore, we extract the time series of Fama-French three factors and risk-free rate from the Kenneth French’s online Data Library. We combine the three datasets together.

The dependent variable is the excess return of individual firms. It is calculated as below in equations (1). We first calculate the log stock return, and then compute the excess stock return by subtracting risk free rate from it.
For the primary explanatory variable, price-to-earning (P/E) ratio, is calculated as the year-end market capitalization value ($M_{it}$) divided by the same year income before interest and taxation ($I_{it-1}$), as shown in equation (2).

\[
\frac{P/E_{it}}{} = \frac{M_{it}}{I_{it-1}}
\]

When analyzing the explanatory power of P/E ratio on stock returns, we also include the FF3 model as the benchmark. These three factors are: Market risk factor (RM - RF), Size risk factor (SMB) and Book-to-Market ratio factor (HML), whereas this multi-factor equilibrium pricing model can be expressed as:

\[
R_{it} = \alpha_i + \beta_{i1}(RM_t - R_f) + \beta_{i2}SMB_t + \beta_{i3}HML_t + e_{it}
\]

- $(RM_t - R_f)$ is the market risk factor which measures the overall market performance.
- SMB is calculated as the return to a portfolio of small capitalization stocks minuses the return to a portfolio of big capitalization stocks.
- HML is the difference between the return to a portfolio of stocks with high ratios of book-to-market value and the return to a portfolio of low book-to-market value stocks.

4 METHODOLOGY

To construct the P/E factor, at the end of June of each year, we sort all stocks in the sample based on their previous year’s annual price-earnings ratio and divide them into 10 portfolios. Each portfolio contains the same proportion of firms of the sample. In particular, portfolio 1 contains the companies with the lowest price-earnings ratio while portfolio 10 contains the companies with the highest price-earnings ratio. Then, starting from July of the same year, we use the classic long-short investment strategy. That is, we hold stocks in portfolio 1 and short-sell stocks in portfolio 10, forming a zero-cost investment strategy. We hold the portfolio for 12 months until the June of the following year. Then we construct a new portfolio ranking in June of the new year and hold the long-short strategy for another year. We repeat the steps until the end of our sample period. We calculate the equally weighted as well as the value-weighted returns for each portfolio. In addition, the average return for the long-short strategy, which is difference in return of portfolio 1 and the portfolio 10, over the sample period is also calculated.

In addition to the simple return series, the risk-adjusted return is also used to measure the performance of the portfolios. Two asset pricing models are used to estimate the risk-adjusted return, namely CAPM and the FF3. This risk-adjusted return series is then alpha when apply the risk models to the actual return data series.
Once we confirm the significance of the long-short investment strategy returns, we then use it as the new pricing factor, namely the PE factor. The new asset pricing equation is then performed as equation (4).

\[ R_{lt} - R_{ft} = \alpha_l + b_{1t}(RM_t - R_{ft}) + b_{12}SMB_t + b_{13}HML_t + b_{14}P/E_t + \epsilon_{lt} \]  

(4)

5 EMPIRICAL RESULTS

Table 1 presents the summary statistics of variables that used in this study, including the number of observation (count), sample average (Mean), standard deviation (Std), minimum (Min), the 25th percentile (25%), median (50%), the 75th percentile (75%), and maximum (Max). As shown in Table 1, there are 124,950 observations for firm return, whose average value is 1.008159, the minimum value is 0.086356, the maximum value is 15.458016, and the standard deviation is 0.134208. In addition, it is apparently, for the PE ratio, the standard deviation is 532.092166, more than 10 times the average. This shows the diversity of companies as the large variation between the maximum value and the minimum value, which are 10284.42 and 0.024265 respective.

<table>
<thead>
<tr>
<th></th>
<th>Equally weighted return</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td><strong>Exc ess return</strong></td>
<td>0.011</td>
</tr>
<tr>
<td><strong>CA PM alp ha</strong></td>
<td>0.005</td>
</tr>
<tr>
<td><strong>t</strong></td>
<td>3.367</td>
</tr>
<tr>
<td><strong>FF 3-alp ha</strong></td>
<td>0.00</td>
</tr>
<tr>
<td><strong>t</strong></td>
<td>3.013</td>
</tr>
</tbody>
</table>

Note: This table presents the summary statistics of variables that used in this study. It includes the number of observation (count), sample average (Mean), standard deviation (Std), minimum (Min), the 25th percentile (25%), median (50%), the 75th percentile (75%), maximum (Max) values of variables.

As shown in Table 2, for each of the three series of returns, we observe a decreasing return trend from portfolio 1 to portfolio 10. For example, the average monthly FF3 alpha showed a downward trend from 0.0028 for portfolio 1 to -0.0020 for portfolio 10, and the t-statistics were
3.013 and -2.158 respectively. There is a same trend in CAPM alpha. The results present that with the increase of P/E ratio, the return of the portfolio shows a decreasing trend. We could expect P/E ratio as a systematic risk factor that would influence asset returns.

### Table 2: Equally weighted returns on P/E ratio sorted portfolios

<table>
<thead>
<tr>
<th>Variable</th>
<th>Count</th>
<th>Mean</th>
<th>Std</th>
<th>Min</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Return</td>
<td>124950</td>
<td>1.00815</td>
<td>0.134208</td>
<td>0.0863</td>
<td>0.9430</td>
<td>1.00000</td>
<td>1.06092</td>
<td>15.4580</td>
</tr>
<tr>
<td>PE ratio</td>
<td>124950</td>
<td>43.3910</td>
<td>532.0921</td>
<td>0.0242</td>
<td>9.0626</td>
<td>14.9310</td>
<td>24.6837</td>
<td>102842.</td>
</tr>
<tr>
<td>Mkt-RF</td>
<td>1152.0</td>
<td>0.008669</td>
<td>0.053417</td>
<td>-</td>
<td>-</td>
<td>0.01060</td>
<td>0.03635</td>
<td>0.38850</td>
</tr>
<tr>
<td>SMB</td>
<td>1152.0</td>
<td>0.00190</td>
<td>0.031745</td>
<td>-</td>
<td>-</td>
<td>0.00090</td>
<td>0.01752</td>
<td>0.36560</td>
</tr>
<tr>
<td>HML</td>
<td>1152.0</td>
<td>0.00354</td>
<td>0.035602</td>
<td>-</td>
<td>-</td>
<td>0.00120</td>
<td>0.01760</td>
<td>0.35610</td>
</tr>
<tr>
<td>RF</td>
<td>1152.0</td>
<td>0.00266</td>
<td>0.002521</td>
<td>-</td>
<td>-</td>
<td>0.0003</td>
<td>0.00420</td>
<td>0.01350</td>
</tr>
</tbody>
</table>

Note: This table describe the equally weighted returns of 10 portfolios sorted by P/E ratio. It also reports the average return of the long-short investment strategy. The top panel reports the original excess return series for the 10 portfolios. And the middle panel presents the risk-adjusted return based on the CAPM model. In addition, the lower panel reports the risk-adjusted return based on the FF3. The t-statistics reported in the second row of each panel. And the last column of this table reports the average return for the long-short investment strategy.

In addition, for the long-short investment strategy, as reported in the last column titled as 1-10, we observe the average return is 0.006228, after adjusting the CAPM model is 0.0073, and after adjusting for the FF3 is 0.0048. Furthermore, the t-statistics for these three sequences are 4.583864, 5.497 and 4.643 respectively. All three t-statistics are greater than 2.32, indicating that the average excess return and the risk-adjusted return based on the CAPM and FF3 are significantly different from 0, at 1% significant level. It is proved that a portfolio constructed based on P/E ratio do generate significant returns.

As shown in table 3, same decreasing trend was observed when sorting the portfolios using the value-weighed return. The average monthly FF3 alpha of portfolio 1 and portfolio 10 are 0.10% and -0.15%, with t-statistics 0.955 and -1.551. There is a same trend in CAPM alpha. For long-short investment strategy, we observe the average return is 0.006228, the risk-adjusted return based on the CAPM model and FF3 are 0.0073 and 0.0048 respectively. The t-statistics for all three sequences are greater than 1.64, indicating that the average excess return and the risk-adjusted return based on the CAPM model and FF3 are significantly different from 0, at 5% significant level. Likewise, it is proved that the long-short investment strategy in table 2 also generates significant returns.
Table 3: Value-weighted returns on P/E ratio sorted portfolios

<table>
<thead>
<tr>
<th>Value-weighted Portfolio</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>1-10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excess return</td>
<td>0.008</td>
<td>0.007</td>
<td>0.008</td>
<td>0.007</td>
<td>0.006</td>
<td>0.006</td>
<td>0.005</td>
<td>0.005</td>
<td>0.006</td>
<td>0.003</td>
<td>0.005</td>
</tr>
<tr>
<td>CA PM alpha</td>
<td>0.003</td>
<td>0.002</td>
<td>0.003</td>
<td>0.002</td>
<td>0.001</td>
<td>0.001</td>
<td>0.000</td>
<td>-0.000</td>
<td>-0.000</td>
<td>0.003</td>
<td>0.006</td>
</tr>
<tr>
<td>t</td>
<td>2.398</td>
<td>2.154</td>
<td>3.475</td>
<td>2.840</td>
<td>1.335</td>
<td>1.855</td>
<td>0.986</td>
<td>-0.174</td>
<td>0.899</td>
<td>-2.948</td>
<td>3.293</td>
</tr>
<tr>
<td>FF3 alpha</td>
<td>0.001</td>
<td>0.000</td>
<td>0.002</td>
<td>0.001</td>
<td>0.000</td>
<td>0.001</td>
<td>0.000</td>
<td>0.002</td>
<td>0.001</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td>t</td>
<td>0.955</td>
<td>0.612</td>
<td>2.372</td>
<td>1.736</td>
<td>0.267</td>
<td>1.553</td>
<td>1.244</td>
<td>0.677</td>
<td>2.730</td>
<td>-1.551</td>
<td>1.703</td>
</tr>
</tbody>
</table>

Note: This table describes the value-weighted returns of 10 portfolios sorted by P/E ratio. And reports the average return of the long-short investment strategy. The top, middle, and lower panels report the original excess return series for the 10 portfolios, the risk-adjusted return based on the CAPM, and the risk-adjusted return based on the FF3, respectively. The t-statistics reported in the second row of each panel. Furthermore, the last column of this table reports the average return for the long-short investment strategy.

All those results confirm that the investment strategy based on the PE factor can generate excess significant return. Our findings confirm with Chu (2016) and Basu (1975, 1977). Chu (2016) directly indicates that the price-earnings ratio is an important indicator for evaluating investment opportunities. According to the results showed in Basu (1975, 1977), low PE ratio portfolios outperform high PE ratio portfolios generating positive and higher risk-adjusted rates of return. And Basu (1977) considered that the price-earnings ratio information is not "fully reflected" in the security price as quickly as the semi-strong form of EMH assumes. The test results of this paper are consistent with this point of view, that is, investing in low P/E ratio stock portfolio did earn higher returns than investing in high P/E ratio stock portfolio. Also, same as Basu (1983) presented, the stock of higher E/P ratio firms did generate above average risk-adjusted return than the stocks of lower E/P ratio firms.

6 P/E PRICING FACTOR

Based on the research results, it is found that P/E ratio can explain the fluctuation of returns, so we consider P/E ratio as a systematic pricing factor and consider its pricing efficiency. We choose to add a new pricing factor, namely, the PE ratio factor. It is calculated as the average return of low P/E ratio firms minus the average returns of high P/E ratio firm. The new PE factor is the difference between the return to a portfolio of Low P/E ratio stocks and the return to a portfolio of high P/E ratio stocks.
We include this P/E ratio as a new risk factor and added it into the traditional FF3. We then conduct the empirical regression and calculate the sensitivity coefficient of the new pricing factor, as shown in equation (4).

The table 4 shows the sensitivity coefficients, beta, of the PE factor. The t-statistics of beta are reported in the second row of each panel. The beta determines how each stock reacts to the PE factor. In Panel A, there is a decreasing beta trend from portfolio 1 to portfolio 10. For example, the beta of PE decrease from 0.5168 in portfolio 1 to -0.4832 in portfolio 10, and the t-statistics were 17.478 and -16.342 respectively. Similarly, In Panel B, it is shown as a decreasing beta trend from portfolio 1 to portfolio 10. The beta of PE decrease from 0.5377 in portfolio 1 to -0.4623 in portfolio 10, and the t-statistics were 28.802 and -24.765 respectively.

For all stocks, the beta is 0.0535 and 0.0795 in Panel A and Panel B, and the t-statistics is 2.747 and 8.729 respectively. The t-statistics of beta is significant, as it exceeds the 1% significant level of 2.32. This suggests that the P/E ratio is a significant systematic factor that would affect stock returns. These results proved that PE risk coefficient is very important to explain the return of portfolio in FF3.

<table>
<thead>
<tr>
<th></th>
<th>Panel A - Equally Weighted</th>
<th></th>
<th>Panel B - Value-Weighted</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FF3</td>
<td>P/E</td>
<td>FF3</td>
<td>P/E</td>
</tr>
<tr>
<td>Mkt Rf</td>
<td>0.9953</td>
<td>0.9996</td>
<td>0.9670</td>
<td>0.9767</td>
</tr>
<tr>
<td>SMB</td>
<td>0.7329</td>
<td>0.7334</td>
<td>0.0081</td>
<td>-0.0107</td>
</tr>
<tr>
<td>HML</td>
<td>0.2849</td>
<td>0.2461</td>
<td>0.0842</td>
<td>-0.0042</td>
</tr>
<tr>
<td>P/E</td>
<td></td>
<td></td>
<td>0.0535</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.0012</td>
<td>0.0009</td>
<td>0.0008</td>
<td>0.0006</td>
</tr>
<tr>
<td>Adj R2</td>
<td>0.950</td>
<td>0.950</td>
<td>0.961</td>
<td>0.965</td>
</tr>
<tr>
<td>N</td>
<td>636</td>
<td>636</td>
<td>636</td>
<td>636</td>
</tr>
</tbody>
</table>

Note: This table shows the sensitivity coefficients of the new pricing factor (PE). Panel A is the 10 equally weighted portfolios, and Panel B is the 10 value-weighted portfolios, those portfolios are all sorted by P/E ratio. In addition, the t-statistics of beta reported in the second row of each panel. Furthermore, the last column of this table reports the beta of PE factor for all stocks.

The table 5 shows the adjusted R-squared results of the FF3 and the new pricing model with PE risk factors. Panel A reports the equally weighted portfolio returns. The adjusted R-squared are both 0.950 based on the FF3 and new pricing model. These figures show the three factors in FF3 can explain 95% changes in stock returns. After adding PE risk factor, the adjusted R-squared did not change. The overall ability to explain returns has not changed. This means that the having PE risk factor does not improve the overall model ability in explaining returns of stocks.

Panel B reports the results based on valued weighted returns, the adjusted R-squared of the new pricing model changes slightly to 0.965, from 0.961 in the standard FF3. These numbers show that the three factors in FF3 can explain 96.1% changes in stock returns. After adding PE risk factor, the overall ability to explain returns has improved slightly, and contribute to explain 96.5% of the returns.
### Table 5: the adjusted R-squared results

#### Panel A - Equally weighted

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_{PE}$</td>
<td>0.516</td>
<td>0.239</td>
<td>0.170</td>
<td>0.143</td>
<td>0.071</td>
<td>0.053</td>
<td>0.004</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.053</td>
</tr>
<tr>
<td>t</td>
<td>8</td>
<td>9</td>
<td>3</td>
<td>4</td>
<td>8</td>
<td>9</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

#### Panel B - Value-weighted

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_{PE}$</td>
<td>0.537</td>
<td>0.198</td>
<td>0.130</td>
<td>0.125</td>
<td>0.055</td>
<td>0.126</td>
<td>0.053</td>
<td>0.060</td>
<td>-</td>
<td>0.462</td>
<td>0.079</td>
</tr>
<tr>
<td>t</td>
<td>7</td>
<td>7</td>
<td>9</td>
<td>6</td>
<td>9</td>
<td>8</td>
<td>3</td>
<td>0</td>
<td>6</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

Note: This table shows the adjusted R-squared results of the FF3 and the new pricing model with PE risk factors. R-squared represents the ability of the risk factors to explain the return on the stock. Panel A is the equally weighted stocks, and Panel B is the value-weighted stocks.

### 7 CONCLUSION

This paper attempts to study the relationship between price-earnings ratio and stock returns. We use the long-short investment strategy, holding stocks with low price-earnings ratio and short-selling stocks with high price-earnings ratio. We then test whether the strategy generate significance earnings. In addition to simple returns, we also calculate the risk-adjusted returns of this long-short strategy, based on the CAPM and FF3. The results show that investment strategy formed on PE ratio generates positive simple returns to investors, as well as positive risk adjusted returns based on CAPM model or FF3. Stocks with lower price-earnings ratio can indeed obtain higher returns than stocks with higher price-earnings ratio.

Then, to test the pricing power of the PE risk factor, the price-earnings ratio is added into the FF3, as a new pricing factor. The results show that PE risk factor has significant positive impact on stock returns, therefore, well explains the return of the investment portfolio. However, the overall model pricing power does not change much. This means the PE risk factor only brings marginal pricing power on asset returns. These findings contribute to the asset pricing literature.

### REFERENCES


Structural Integration and Optimization of Human Resource Management Under the Background of Artificial Intelligence

Linqiang Zhang\textsuperscript{1,}\textsuperscript{a}, Chang Wang\textsuperscript{1}

m15170456008@163.com\textsuperscript{a}

Nanchang Institute of Technology, Nanchang, Jiangxi Province, 330044, China\textsuperscript{1}

Abstract: In order to improve human resource management models from the basis of artificial intelligence, this project conducts research on the use of big data and artificial intelligence in human resource management, in the context of the rapid development of intellectual technology. Using Company G as an example, the problem of human resource management and why the above problems occur is studied through a combination of structural analysis, qualitative methods and more. After a lot of analysis, we can get the results of three things that happen in the company's human resources management: recruitment, human resources planning and performance management.

Keywords: Artificial Intelligence, Human Resource Management, Structural Integration and Optimization, Big Data, Qualitative Methods, Management Issues, Performance Management, Recruitment.

1 INTRODUCTION

Artificial Intelligence technology has brought new growth points over time. Recruiting skills continue to improve and change in the direction of digitalization. Optimizing the organization's structure and seeing the perception of employees show the steps of enterprise reform and change. In this process, if the internal human management of the company is difficult to follow the pace of development and meet the new goals of the enterprise, it will be difficult for us to be profitable and easy. technology development. In reality, it is still affected by problems such as poor screening, poor job and job comparison, and superficiality of traditional recruitment technology [1].

2 METHOD

2.1 Definition of Basic Concepts

For business, human resources are complex and require us to plan and coordinate. Therefore, human resource management is to determine the direction of the company's future goals, to achieve sustainable development, and to provide opportunities for all employees to use their strength. Therefore, this type of management is important for businesses [2-3].
The main concept of artificial intelligence is that the machine is able to learn and be intelligent. This learning ability refers to the ability to practice an action that can be explained. The term "artificial intelligence" was first coined by John at a conference at Dartmouth College [4]. Artificial intelligence is a complex group that encompasses many things. It is a science and technology that people hope to achieve many tasks by controlling machines such as knowledge, analysis, and work. His intelligence has three aspects: knowledge, imagination, and calculation [5].

2.2 Human Resource Structure Analysis of the Company A

In 2018, the number of employees in Company A has reached 1210. At present, Company A needs to adopt the team management mode, that is, employee area and products. Only in this way, the dual leadership in business can realize the sharing of human products and services between departments, maximize the utilization of equipment and other resources and enterprise resources according to different human resources strategies. The composition of human resources of Company A is shown in Table 1 [6-7].

<table>
<thead>
<tr>
<th>Enterprise employee category</th>
<th>Number of employees (person)</th>
<th>Proportion (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enterprise general staff</td>
<td>912</td>
<td>75.4</td>
</tr>
<tr>
<td>General management staff of the enterprise</td>
<td>188</td>
<td>15.5</td>
</tr>
<tr>
<td>Skilled staff</td>
<td>47</td>
<td>3.9</td>
</tr>
<tr>
<td>Senior management staff</td>
<td>63</td>
<td>5.2</td>
</tr>
</tbody>
</table>

Table 1 Composition of human resources of Company A

At present, employees of Company A generally have college education or above, managers have bachelor's education, and technicians have master's education. Most people have doctorates. See Table 2 for details [8].

<table>
<thead>
<tr>
<th>Employee category</th>
<th>Educational background below junior college (number and proportion)</th>
<th>College degree (number and proportion)</th>
<th>Bachelor degree (number and proportion)</th>
<th>Master degree (number and proportion)</th>
<th>Doctor degree (number)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Production line staff</td>
<td>603/49.8%</td>
<td>212/17.5%</td>
<td>111/9.2%</td>
<td>0/0</td>
<td>0</td>
</tr>
<tr>
<td>General management staff</td>
<td>0/0</td>
<td>76/6.3%</td>
<td>104/8.6%</td>
<td>12/1%</td>
<td>0</td>
</tr>
<tr>
<td>Skilled staff</td>
<td>0/0</td>
<td>0/0</td>
<td>11/0.9%</td>
<td>20/1.7%</td>
<td>3</td>
</tr>
<tr>
<td>Senior management staff</td>
<td>0/0</td>
<td>0/0</td>
<td>40/3.3%</td>
<td>12/1%</td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td>603/49.8%</td>
<td>288/23.8%</td>
<td>266/22%</td>
<td>42/3.7%</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 2 Human resource structure of Company A
2.3 Problems and Causes of Company as Recruitment

The recruitment method of Company A is too traditional. Nowadays, the advent of information technology and the information age can provide more employment opportunities. However, the company's recruitment method is still traditional, many shortcomings. First, many candidates do not accept return letters [9]. Second, it is difficult to work with many different documents. Third, despite high labor costs and low productivity [10].

2.4 Problems and Causes of Human Resource Development Planning of Company A

Company A's business grew rapidly and its profit doubled between 2017 and 2019. However, the rapid development of the economy brought problems. The shortage of human resources is increasing [11]. The workforce is now under pressure due to delays in recruitment and training. To improve jobs and create new jobs, it is necessary to increase the number of employees. These difficulties then hinder the development of A, the internal knowledge of traditional production, the material and the need for intellectual skills reduce the demand for work, and the work patterns of skilled workers are also changing. If artificial intelligence destroys the human brain, it will change many jobs and in the long run, the work of traditional business will decrease. Therefore, the best work will be promoted for the talented people [12-13].

3 RESULTS AND ANALYSIS

Now, even though Company A is following the current development, there are still problems that need to be solved.

Company A is still at a low level in human resource management, especially in labor management, which is still common. The use of traditional methods will cause us to face the problems of the next period of inflation [14]. In order to avoid this problem, we should start from this aspect now, apply big data to performance management, realize the integration and optimization of resources, and improve work efficiency, including the following aspects:

First, the company can create a quality management system for this problem, access data, and use the system as a communication platform for matrix management problems, which yes, regional leaders and business leaders. At the same time, the difference is that managers in different industries manage their own work and make different kinds of employees. At the same time, employees communicate directly with the work team to improve communication and provide quick feedback and corrections [15].

Second, the company can establish a management system for this problem and optimize the system configuration in the system. This system only uses big data to collect all the digital information of the employees of the company, establish a database, use appropriate models to analyze employees' work ability and work style behavior, and turn the results of data into data that can measure employees' performance. The establishment of the system takes into account the principles of comprehensiveness, comprehensiveness and scientificity, and solves the problem of imperfect indicators of the company. The assessment results are also fair.

Third, the company can establish a performance management system for this problem, analyze the data of employees, improve the evaluation efficiency, save human resources and reduce the
pressure of human resources departments on talent selection, which allows the human resources team to focus on resource allocation and planning [16].

To improve the company's internal talent, it means that the company can use the existing human resources more effectively, reducing the inefficiency of recruiting more employees frequency, and save a lot of money on human costs. During the development of the human resources system of Company A, a lot of information was collected, the information was processed and distributed. By classifying, processing and summarizing these information data through artificial intelligence technology, human resources departments can have a clearer understanding of talent training and development, making it more reasonable and scientific [17].

Company A can use a combination of big data and artificial intelligence to process additional data and use it to evaluate employee performance. Although based on the results of these evaluations, people and jobs can be improved, the evaluation of employee integrity, the identification of potential employees, the employee turnover, etc. can be used in conjunction with real events. This can have a positive impact on the company's employees in terms of retention, training and development [18].

This plan will automatically generate the results of the company's quarterly evaluation for Q1-Q4 and provide suggestions for the development of employees at the same time. Through this application, all employees can understand the work feedback at any stage in time and use the improvement plan provided by the human resource management system. nature to develop training plans and subsequent development. Figure 1 [19-20].

4 CONCLUSION

This article uses a combination of analysis and performance and various methods to explore the problems and make people management, focusing on the use of human skills resources in the business, set the specific process of application in the business, and take the company as an example. Accordingly, this document rationalizes the problems of human resource management, helps the managers of enterprises to ensure the competitiveness of big data, improve management resources in business, promote economic development, and promote development. The skills of the units are a competitive advantage of the company.

Acknowledgement: Science and Technology Research Project of 2021 Education Department of Jiangxi Province
Research on the challenges and opportunities of human resource management in the era of artificial intelligence no. GJJ212119

REFERENCES


Novel Blockchain-Based Privacy Protection for Smart Home

Lifeng Pan¹, Shuguang Zhou²

lifengpan@ln.hk¹, 3200104014@zju.edu.cn²

Department of Computing and Decision Sciences, Lingnan University, 8 Castle Peak Road, Hongkong, China¹
Department of Mathematical Sciences, Zhejiang University, No. 866, Yuhangtang Road, Hangzhou, China²

Abstract: Smart home data management systems are starting to be put into use in many smart home companies. To protect the private data of users in these smart home data management systems, the paper proposes Smart Home Guard, a blockchain-based system for recording smart home data operations, which creates smart contracts deployed on the Ethereum private blockchain based on Solidity. Users can customize data access policies, authorize and monitor queries on private data. By taking advantage of the immutability of the blockchain, the query records of user data are monitored and traced back to prevent user data from being stolen or maliciously tampered with. Smart Home Guard is designed to help regulators track the use of smart home data. At the same time, it allows arbitrators to verify the authenticity of evidence in the event of a privacy breach dispute.

Keywords: Blockchain, Smart Home, User Privacy, Data Protection.

1 INTRODUCTION

With the extensive use of smart home devices, their imperfect security protection capability is highly susceptible to causing large-scale leakage of user privacy data. As shown in Figure 1, the device security analysis of smart home is mainly lies in the perception layer, network layer and application layer [5]. The purpose of this paper is to design a blockchain-based smart home system to record user data operation - Smart Home Guard, by using the tamper-evident nature of blockchain to query the user data. By using the immutability blockchain, the records are monitored and traced to prevent user data from being maliciously tampered and stolen, and to achieve the protection of user data security in the information interaction stage, i.e., the network layer.

According to access and management authorities, blockchain is divided into three categories: public blockchain, consortium blockchain and private blockchain. In this paper, private blockchain is used to build Smart Home Guard for the following main reasons:

1) Smart home system does not need too many nodes, and private chain retains the advantages of node communication while avoiding network attacks brought by too many communication nodes to a great extent.
2) The data on the public chain is highly transparent and not suitable for privacy protection, while the data on the private chain is highly private and users can even restrict and authorize read-and-write access by modifying the smart contract.

Figure 1: Smart home security analysis focus.

Compared with traditional distributed data management systems, Smart Home Guard has the following advantages.

1) Immutability: In the security solutions of traditional information system, security relies on layers of guarded access control. Smart Home Guard uses blockchain technology to record transactions in a database that can be accessed by anyone, but by clever design and supplemented by cryptography and consensus mechanisms, if an attacker modifies a data, he must change all subsequent data.

2) Heterogeneous multi-activity: Each system participant in Smart Home Guard is an off-site multi-activity node, which is a multi-activity system essentially. If a node is controlled by hackers or encounters network problems, hardware failures, software errors, it will not affect the system and other participating nodes.

Smart Home Guard, the data management system server is connected to the blockchain server, calls the smart contract interface, and adds records and queries information through the Ether client. When a user adds data, an Ethernet Virtual Machine (EVM) generates a new block. Then the Ethernet server adds it to the blockchain through a "mining" operation. Querying data does not generate new blocks. Therefore, the blockchain server in this paper stores information about the operations performed by visitors, such as additions, deletions, updates, and queries. These records are permanent and cannot be changed.

2 Relevant Theoretical Foundations

The application of blockchain technology in IoT is one of the research hotspots in recent years. Xiubo Liang et al. [6] analyzed data storage security, privacy security, data access security and
data sharing security four aspects, the blockchain technology facing data security problems and related technology solutions. Lihua Song et al. [14] proposed a model to improve the access control security of the IoT, which was based on zero-knowledge proof and smart contract technology in the blockchain. Liu Tao et al. [7] proposed a service architecture platform of blockchain technology for the terminal connection of IoT. D.M Sheeba et al. [15] believes that the use of blockchain in the IoT environment provides the flexibility to handle large amounts of data in a secure manner between IoT applications and consumers.

In terms of data storage and privacy protection, Zixiong Zhao et al. [18] proposed that the peer-to-peer nature of some nodes can be sacrificed to build a polycentric rather than peer-to-peer blockchain, where network nodes can be properly managed in a hierarchical manner. Nazar Waheed et al. [17] presented the current solutions to IoT security and privacy by utilizing machine learning algorithms, blockchain techniques, and the integration of both.

In summary, many researchers have applied blockchain technology to IoT, which proved that blockchain technology is applicable to the field of IoT. Then as an area of the Internet of Things, the smart home also applies. However, most of the research are focused on the field of data sharing, which requires significant changes to existing information systems, rather than focusing on supervision and evidence retention. Rizwan Majeed et al. [11] presented a novel idea of a smart home that uses a machine learning algorithm (Support Vector Machine) for intelligent decision making, it also uses blockchain technology to ensure identification and authentication of the IoT devices. Jung Hyun Ryu et al. [12] presented a digital forensics framework for the IoT environment based on the blockchain technology. By using blockchain technology, the integrity of the data to be analyzed has been ensured and security has been strengthened, and the preservation of integrity is more reliable by a decentralized method of integrity preservation.

3 DATA COLLECTION

To maximize deployment convenience, the data operation record collection design of Smart Home Guard is decoupled from the smart home data management system as far as possible. Currently, Smart Home Guard implements two types of data: database monitoring and private API. [13] Table 1 lists the types of data currently collected by Smart Home Guard.

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Data Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>time</td>
<td>access time</td>
<td>timestamp</td>
</tr>
<tr>
<td>ip</td>
<td>access ip</td>
<td>string</td>
</tr>
<tr>
<td>id</td>
<td>name of the data visitor</td>
<td>string</td>
</tr>
<tr>
<td>op_type</td>
<td>operation type</td>
<td>string</td>
</tr>
<tr>
<td>data_type</td>
<td>type of user data accessed</td>
<td>string</td>
</tr>
<tr>
<td>devices</td>
<td>devices in this operation</td>
<td>string</td>
</tr>
</tbody>
</table>
3.1 Data Base Monitoring

Currently, most smart home data management systems have built-in logging capabilities to record important information such as access time, visitor ID, and visitor IP. To minimize the deployment cost, Smart Home Guard has implemented a set of database middleware [2] which is used to track the log records of smart home data management systems and extract smart home user data access information from database requests based on pre-configuration. Currently, Smart Home Guard implements middleware for MySQL only, and its structure and workflow are shown in Figure 2.

![Figure 2: Database Monitoring Architecture.](image)

The middleware implements a standard MySQL database interface, running on port 3306 by default, for receiving SQL data from the back-end program of the smart home data management system and then hand it over to the analyzer and database server. The parser is used to extract the smart home user data access information from the SQL statements, as shown in Figure 3. After finishing the information extraction, the parser processes the data to generate structured data and uploads the structured data to the blockchain through JSON RPC. [3,16]

![Figure 3: Analyzer Architecture.](image)

3.2 Private API

Ideally, smart home data management systems mostly use common databases such as MySQL, Oracle, and MSSQL [8], but this is not the case, and many of them do not use these. In addition,
some smart home data management systems even have proprietary designs, and information is difficult to monitor through databases. For this situation, this paper designs and implements a private API that developers can use to receive data reports from smart home data management systems. The workflow of this interface is shown in Figure 4. When the user reads and writes the smart home data management system, the system back-end program performs a series of database operations, and after all the database operations are completed, the system back-end program calls the private API of Smart Home Guard to report the data to the blockchain.

4 SYSTEM DESIGN

In this section, we will detail the architecture design and implementation of Smart Home Guard, which uses blockchain technology to provide complete records of user access and operation data. The traceability and immutability of blockchain can be used to establish a credible third party while providing effective data support and supervision for regulators and the judiciary. In addition, with the function of monitoring data-level suspicious users, Smart Home Guard can help regulators detect internal data thieves and large amounts of unauthorized data downloads by external attackers quickly.

4.1 System Architecture

The architecture of Smart Home Guard is shown in Figure 5, which includes three components: Blockchain Infrastructure, Agent, and User Platform. The blockchain infrastructure is responsible for implementing all blockchain-related functions. As the core module, it includes data storage and permission control, and is the core module. Agent is deployed on the smart home data management system server, which used to monitor the operations of the smart home data management system and upload the relevant information to the blockchain module. For management, the user platform is responsible for monitoring suspicious users. Authorized users of the regulator can query the operation records of various smart home data management systems through this platform.

Figure 4: Private API Process.
4.2 Blockchain Infrastructure

The private blockchain of Smart Home Guard is based on Ethernet and deployed in a virtual private network, and the smart contract is deployed on the private blockchain. The blockchain infrastructure is used to complete blockchain related operations, such as logging, querying, adding, deleting, updating, checking, etc. through its own smart contract and JSON RPC interface. The blockchain infrastructure consists of two main parts, private blockchain and smart contracts. The latter is deployed on the private blockchain. It implements several interfaces which are used for reading and writing data and verifying permissions. As shown in Table 2, to facilitate the backward compatibility of smart contracts, variable names are not defined according to the actual field meanings during the contract development process. Also, several data fields are reserved for data storage.

The smart contract implements the following functions as interfaces.

1) Uint2Str, which converts unsigned integer data to strings.
2) StrConcat, stitching 4 strings into one long string.
3) TwoStrConcat, which splices 2 strings into a single string.
4) Convert, which splices 4 input parameters into a long string, and then converts it to Byte32 type.

<table>
<thead>
<tr>
<th>Name</th>
<th>Name in Smart Contract</th>
<th>Used in Search Record</th>
</tr>
</thead>
<tbody>
<tr>
<td>time</td>
<td>timestamp</td>
<td>Yes</td>
</tr>
<tr>
<td>ip</td>
<td>keyArg1</td>
<td>Yes</td>
</tr>
<tr>
<td>id</td>
<td>keyArg2</td>
<td>Yes</td>
</tr>
<tr>
<td>N/A</td>
<td>keyArg3</td>
<td>yes (reserved, null by default)</td>
</tr>
<tr>
<td>devices</td>
<td>infoArg1</td>
<td>No</td>
</tr>
<tr>
<td>op_type</td>
<td>infoArg2</td>
<td>No</td>
</tr>
<tr>
<td>data_type</td>
<td>infoArg3</td>
<td>No</td>
</tr>
<tr>
<td>N/A</td>
<td>infoArg4</td>
<td>No</td>
</tr>
</tbody>
</table>
The smart contract implements the interface functions for manipulating relevant information of blockchain, including adding, deleting, updating and querying. The specific code is as follows:

Input: timestamp, keyArg1, keyArg2, keyArg3, infoArg1, infoArg2, infoArg3, infoArg4;

Function AddRecord Input Arguments

- put timestamp, keyArg1, keyArg2 and keyArg3 into a structure keys;
- put the new keys into the global array of structure named keysArray, to store the index information;
- \( A = Convert(\text{timestamp}, \text{keyArg1}, \text{keyArg2}, \text{keyArg3}); \)
- calculate the hash value of \( A \) and assign it to key as index;
- put infoArg1, infoArg2, infoArg3, infoArg4 and the address of user into a structure info;
- put the key and info into global map named records, use key as index;

In addition, the contract implements an interface that allows users to retrieve, add, delete, modify and search records of operations from the blockchain, the specific code is as follows:

Input: timestamp, keyArg1, keyArg2, keyArg3;

Function GetRecord Input Arguments

- \( A = Convert(\text{timestamp}, \text{keyArg1}, \text{keyArg2}, \text{keyArg3}); \)
- calculate the hash value of \( A \) and assign it to key as index;
- get the values corresponding to key from global map records;
- return values;

### 4.3 Agent

Agent runs on the smart home data management system server, and each system corresponds to an Agent, which collects data from it and reports to the blockchain. The operation and maintenance engineer can modify the configuration file and then choose the operating mode of the Agent (database monitoring or private API). The workflow of the Agent is shown in Figure 6.

![Figure 6: Workflow of Agent.](image-url)
The blockchain communication module interacts with the blockchain based on JSON RPC. After encoding, analyzing and recording, the module sends the obtained log data to the blockchain. The database module receives SQL statements from the smart home data management system and then forwards them to the parser and database server of the smart home data management system. The private API module implements a microservice API locally to process and clean the data manipulation interface provided by the smart home data management system. It can also provide structured data to the analyzer. The analyzer extracts the required information from the data passed by database module and private API module, and invokes the blockchain communication module to complete the relevant operations. The analyzer maintains a message queue for asynchronous operation to prevent the Agent from interfering with the operation of the smart home data management system.

4.4 User Platform

The user platform is an interaction platform provided by Smart Home Guard to the regulator. The regulator can query the data creation, modification and other operation records of all smart home data management systems connected to Smart Home Guard through this platform, while the platform sends alerts to the regulator for unauthorized bulk data access through the data recorded in the blockchain. The user platform includes 3 components: the Web front-end server, the Web back-end server, and the security audit server, as shown in Figure 7.

In the user platform, the security audit server obtains data from the blockchain infrastructure and performs security analysis. The web front-end server is based on the Apache HTTP Server [4]. Smart Home Guard provides visitors with a user-friendly and beautiful React-based human-computer interface. The web back-end server is implemented on top of Sanic, a high-performance web framework built on Python. The web backend server is mainly responsible for the actual data processing of the user platform, including operations such as obtaining smart home data operation records from the blockchain infrastructure, obtaining security alerts from the security audit server, and providing information to users through the Web frontend server.
Currently, the main threat to the smart home data management system is the theft of data by attackers, including a variety of attack types. Except for external penetration, the attackers may also be internal malicious users. Most of the traditional smart home data management systems are deployed on internal networks, which are difficult to access by external users, so the risks mainly come from internal. However, with the further development of the Internet, more and more smart home companies are providing APP remote services, which allow users to check their smart devices and operation records through the Internet. This service is usually realized through secondary development of existing smart home data management systems by smart home organizations, which enables external attackers to use vulnerabilities to attack the smart home data management system, which undoubtedly increases the risk of data. Smart Home Guard mainly implements security monitoring for batch operations (data downloading, data modification and so on) identified as authorized users by the smart home data management systems. This is mainly to respond to internal attackers and vulnerabilities like horizontal privilege escalation and vertical privilege escalation, which are difficult to handle by existing Web security defense such as Web Application Firewall (WAF) and Intrusion Detection System (IDS) (BACUDIO 2011, LIAO 2013).

5.1 Design and Implementation

Smart Home Guard's security auditing is implemented based on rules for detecting smart home data access records in the blockchain, with a core approach of raising alerts when abnormal batch data operations are detected. The security audit server is a part of the Smart Home Guard user platform. It includes a blockchain communication module for fetching data from the blockchain infrastructure, an analyzer for analyzing recent smart home data operation records, a lightweight database for storing alert information, and a microservice API module for receiving user calls and monitoring alerts, as shown in Figure 8.

![Figure 8: Security Audit Server.](image)

5.2 Rule Definition

As for detecting rule-based malicious behavior, the most important thing is to institute reasonable and effective rules. In this paper, we extract following features and set thresholds in the data recorded by Smart Home Guard's blockchain infrastructure:

1) $S_{user}$: The set of data operations for a specified user.

2) $S_{devices}$: The set of operations performed on a specified device in a day.

3) $S_{I}$: Set of a specific user’s interval time between two consecutive data operations.
4) \(l_{time}\): Specifies the threshold value of the interval between two consecutive operations by the user, which is determined by the regulator, i.e., the enterprise.

5) \(l_{count}\): A threshold value specifying the number of device operations, determined by the regulator, i.e., the enterprise, and must be smaller than the buffer size.

6) \(S_{optype}\): The set of all operation types of a certain operator.

This paper uses these characteristics to define the following rules. The analyzer sends an alert when user’s behavior meets any of them.

1) Fast and continuous operations. Based on existing experience and historical vulnerability reports, attackers usually use fast and continuous operations when they try to steal data. Therefore, when \(\Delta_{fi}\) is the \(i\)th item in \(S_{\Delta f}\), we define the following rule:

\[
\forall \Delta t_i \in S_{\Delta f}, \Delta t_i < l_{time}
\]  

(1)

2) Similar operation interval. Experienced attackers would try to hide themselves. They may perform a data operation every few seconds, which means that there tends to be a regular interval between the two operations. Therefore, we define the following rule:

When \(\Delta_{fi}\) is the \(i\)th item in \(S_{\Delta f}\), if \(i\) meet

\[
0 < i < |S_{\Delta f}|
\]

(2)

We have

\[
\forall \Delta t_i \in S_{\Delta f}, |\Delta t_i - \Delta t_{i+1}| < l_{time}
\]

(3)

3) Too many data operations in one day. An attacker who aims to obtain data in batches through illegal means has a much larger amount of data access than normal users, so we designed a threshold. The rule defined with this threshold are shown as following. Users can set this threshold value according to the scale of the smart home organization.

\[
|S_{device}| < l_{count}
\]  

(4)

4) Same type of operation. Experienced attackers typically use scripts to perform slow batch processing of data sets during unobtrusive hours such as midnight. These scripts usually do not perform operations outside of the task and work in a completely different mode from the average user. If a smart home data management system user performs the same type of operation continuously, then the operator is probably a malicious user. Therefore, we define the following rule:
If $s$ meet

$$s \in S_{optype}$$

We have

$$\forall p \in S_{optype}, s = p$$

6 CONCLUSION AND FUTURE WORK

This paper aims to design a blockchain-based smart home user privacy protection system - Smart Home Guard, which will be applicable to the existing data systems at all ends of the smart home field. Smart Home Guard can help users and enterprises realize unmodifiable monitoring of data access, facilitate users and enterprises to view data usage records, and serve as a strong evidence when data leakage is found. Smart Home Guard has proven to be effective in meeting design requirements and providing regulators with a low-cost solution for monitoring smart home data usage. However, the current version of Smart Home Guard has some problems: firstly, it is based on the private blockchain of Ether, and its performance and security are limited by Ether. Secondly, the tracking of smart home device operation is not accurate enough. For example, in order to be more compatible with the existing smart home data management system, the existing data collection technology cannot distinguish between the same-named devices in the same system. Finally, the security audit mainly guards against bulk data operations by malicious attackers. Because of its over-reliance on rules, it is unable to find attackers with more ambiguous behavioral characteristics. The next step will be developing a new blockchain infrastructure for Smart Home Guard and using new algorithms, such as Byzantine Fault Tolerance (BFT) \[10\], to enhance security.

Acknowledgements: Thank Dr. Liu Xin of Lanzhou University for his help in this paper.

REFERENCES

The Influence of Institutional Investor Shareholding on the Innovation Level of Listed Companies

Qingyuxi Yang¹, Hangqin Xiang²,*

1397305313@qq.com¹,
*Corresponding author: hxia5685@uni.sydney.edu.au, phone18302359600,²

Business School, Sichuan University, Chengdu, Sichuan, China¹
Faculty of Engineering, The University of Sydney, Sydney, NEW, Australia²

Abstract: This paper aims to study the interaction between institutional investor ownership and innovation level of listed enterprises in China. Using Python crawling and manually collected data of listed companies as samples, the innovation level is measured by two dimensions of innovation input and innovation output, and the innovation output is divided into high and low standard innovation. The results show that the shareholding ratio of institutional investors has a negative effect on the proportion of R&D investment and R&D personnel of listed companies. Furthermore, institutional ownership has a negative correlation with high standard innovation, but no correlation with low standard innovation.

Keywords: Institutional Investors, Technological Innovation Capability, R&D, Patent License.

1 INTRODUCTION

Innovation plays a key role in the survival and development of enterprises. The level of innovation is affected by a series of factors, the institutional shareholding factor is more prominent. At present, the academic circles have not formed a consensus on the connection between the two, which can be divided into three schools: effective supervision theory, myopic theory and irrelevant theory.

According to the effective supervision theory, institutional investors have a better understanding of the internal and external conditions of enterprises, have a positive direction for the long-term returns of innovation activities of enterprises, and will urge enterprises to increase their R&D investment to increase the return on investment of institutions [1]. Long-term R&D projects of enterprises are beneficial to enhance the competitiveness of enterprises. In order to ensure future profitability, organizations will pay attention to the stability of the rate of return of the company's projects [2]. According to Wahala & MC Connell, one of the main reasons for the increase in corporate investment spending is institutional ownership, and the level of R&D spending increases as institutional ownership increases [3]. Based on the agency cost theory and the system characteristics of China's financial market, Henry found that with the increase of the share quota of institutional investors, the shortcomings of innovation input of enterprises could be better avoided [4]. Due to the rights obtained by holding shares, institutions are more willing to participate in the daily management of the company, so as to exert influence...
on strategy formulation and investment project selection\(^5\). Hu Yuming and Fan Haifeng recognized the convergence of interests between institutions and enterprises, and believed that they had common goals and deep connections, including interests and reputation\(^6\). He Mingqin found that most listed companies with certain institutional investors have a large amount of expenditure on innovation projects\(^7\).

Myopic theory, in contrast, argues that institutions are more keen on short-term returns because of market evaluations and the pressure to report regularly on their investments\(^8\). Mitra also believes that the reason why institutions are averse to high-risk long-term investment may be the impact of performance assessment faced by institutions. The management will reduce innovation activities to retain institutional shares and prevent the impairment of the company's stock\(^9\). Ren Haiyun also concluded that institutional investors have similar behavior of "seeking advantages and avoiding disadvantages" based on the assessment methods in the institutional industry\(^10\). Non-independent institutional investors believe that supervision will increase their business costs\(^11\). Based on information asymmetry, institutions may think they are at an information disadvantage and hold a pessimistic attitude toward long-term investment activities that require a large amount of information\(^12\). Chinese scholars believe that Chinese institutional investors are still in the initial stage of development, and the governance of participating enterprises is still in a wait-and-see state\(^13\). An Tongliang and Qi Jiebin took a different approach, which was different from the linear conclusion of previous studies, and found that different institutional investment shareholding ratio had different impacts on the innovation ability of enterprises. There was a critical point, namely the so-called threshold, and only when the shareholding ratio reached a certain level would it have a positive impact on the innovation ability of enterprises\(^14\).

Unrelated theory thinks that institutional holding has no correlation with innovation ability of listed company in our country\(^15\).

To sum up, the possible reason is that the internal organization is a diversified whole, and different types of institutional investors have different degrees of participation in enterprise management. The countervailing of effective supervision theory and myopic theory and the different degrees of countervailing eventually lead to three different views, which need further study.

2 THEORETICAL ANALYSIS AND HYPOTHESIS

Institutional investors have natural advantages, such as professional team and capital scale, and it is easier to obtain internal and external information of the enterprise to judge the long-term value of the enterprise, so as to support the R&D investment of the management. However, due to reputation considerations and performance competition under market expectations, some institutions tend to obtain short-term benefits through portfolio adjustment and avoid high-risk long-term capital injection projects as much as possible. Therefore, the following alternative hypotheses are put forward:

H1a: The shareholding of institutional investors is positively correlated with the R&D investment intensity of enterprises.
H1b: The shareholding of institutional investors is not correlated or even negatively correlated with the R&D investment intensity of enterprises.

There are three types of authorized patents in China -- invention types, utility models and design types. The patents with a high level of innovation and technology belong to the invention patents, and the patents that only improve the appearance or performance of the products are called design patents and utility model patents. From the perspective of R&D difficulty and application difficulty, the invention patent belongs to the high standard patent, and the other two types belong to the low standard patent. Information asymmetry makes customers believe that the products of enterprises with more authorized patents will be better and more competitive. Enterprises are willing to advertise the number of national patents as their advantages, so as to attract the investment of institutional investors. In addition, China implements the innovation-driven development strategy, the government encourages the innovative behavior of enterprises, and spends funds on the application of patents, which reduces the innovation cost of enterprises and relieves the capital pressure of enterprises. However, some institutions prefer the high rate of return and long-term benefits brought by high-standard patents. However, the scarcity of resources and limited capital also determine that more investment in high-standard innovation leads to less attention on low-standard innovation. To sum up, the following hypotheses are proposed:

H2: Institutional investor ownership has a reverse effect on high-standard innovation, but has no obvious effect on low-standard innovation.

3 STUDY DESIGN

3.1 Sample Selection

In this paper, A-share listed companies with complete R&D investment and number of patent grants for ten consecutive years from 2010 to 2019 were selected as the original research samples. Part of the data were from CSMAR, Juchao Information Network and publicly disclosed annual reports of listed companies. Supplementary data were obtained through Python crawling and manual collection.

3.2 Variable Design

1) Explained variable: innovation level

Innovation investment: the ratio of R&D investment to operating revenue & the ratio of researchers to total employees.

Innovation output: high standard patents and low standard patents.

2) Explanatory variable: Shareholding ratio of institutional investors.

3) Control variables: company Size (Size), financial leverage (LEV), company Growth (Indep), company performance (ROE), ownership concentration (TOP10ratio), company value (TobinQ), Dual.

Specific variable names and definitions are shown in Table 1.
TABLE I.  TABLE OF VARIABLES AND INDICATORS

<table>
<thead>
<tr>
<th>Variable type</th>
<th>Variable symbol</th>
<th>Variable definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explained variable</td>
<td>IRatio</td>
<td>The proportion of shares held by institutional investors to the total share capital of listed companies</td>
</tr>
<tr>
<td>Explained variable</td>
<td>RDSpendSumRatio</td>
<td>The proportion of the total amount of R&amp;D investment in the company's main business income at the end of the year</td>
</tr>
<tr>
<td>Explained variable</td>
<td>RDPerson Ratio</td>
<td>The proportion of R&amp;D personnel in all employees of the enterprise</td>
</tr>
<tr>
<td>Explained variable</td>
<td>Invia</td>
<td>High standard innovation</td>
</tr>
<tr>
<td>Explained variable</td>
<td>zpata</td>
<td>Low standard innovation</td>
</tr>
<tr>
<td>Explanatory variables</td>
<td>Size</td>
<td>The natural log of total assets at year-end</td>
</tr>
<tr>
<td>Explanatory variables</td>
<td>Lev</td>
<td>Asset-liability ratio = total liabilities/total assets at year-end</td>
</tr>
<tr>
<td>Explanatory variables</td>
<td>Growth</td>
<td>Growth rate of operating income</td>
</tr>
<tr>
<td>Explanatory variables</td>
<td>Indep</td>
<td>The ratio of independent directors to the board of directors</td>
</tr>
<tr>
<td>Explanatory variables</td>
<td>ROE</td>
<td>Return on equity</td>
</tr>
<tr>
<td>Control variables</td>
<td>TOP10ratio</td>
<td>The proportion of total shares held by the top 10 shareholders</td>
</tr>
<tr>
<td>Control variables</td>
<td>TobinQ</td>
<td>Measured by Tobin Q</td>
</tr>
<tr>
<td>Control variables</td>
<td>Dual</td>
<td>Whether the chairman of the board and the general manager concurrently</td>
</tr>
</tbody>
</table>

3.3 Model Building

This study the following multiple regression model was constructed, in which F enterprise technology innovation ability, RDs of r&d spending accounted for the enterprise, RDp for enterprise research and development personnel accounted for, Invia high standards for the enterprise innovation, the patent zpata low standards for the enterprise innovation patent license number, the Control for the set of Control variables in the table above, epsilon for random disturbance.

To test the correlation between the shareholding ratio of institutional investors and the innovation level of enterprises:

\[ RDs = \alpha + \alpha_1 IH + \alpha_2 Control + \varepsilon \]  
\[ RDp = \beta + \beta_1 IH + \beta_2 Control + \varepsilon \]  
\[ Invia = \delta + \delta_1 IH + \delta_2 Control + \varepsilon \]  
\[ zpata = \eta + \eta_1 IH + \eta_2 Control + \varepsilon \]

To sum up, the influence of overall shareholding of institutional investors on technological innovation capability of enterprises is measured as follows:
\[ F = \alpha R_Ds + \beta R_Dp + \delta \text{Invia} + \eta \text{zpata} + \varepsilon \] (5)

4 EMPIRICAL TEST AND RESULT ANALYSIS

Table 2 shows the descriptive statistics of each variable in this paper.

The average ratio of R&D intensity is 4.85%, which is lower than the internationally recognized 5%. The minimum value is 0, and the maximum value is 88.56%, indicating that there are significant differences between the R&D investment intensity of A-share listed enterprises. Some enterprises attach importance to the level of R&D investment, but the overall level of R&D still needs to be improved. The average shareholding ratio of institutions is 35.14%, the lowest is 0, the highest is 98.66%, and the standard deviation is 0.2374, indicating that there is a certain difference in the share ratio of listed companies owned by institutional investors. The average number of patented inventions authorized by listed companies is 12, with a standard deviation of 94.6523.

<table>
<thead>
<tr>
<th>variable</th>
<th>N</th>
<th>mean</th>
<th>sd</th>
<th>min</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>RDPerson Ratio</td>
<td>10,579</td>
<td>0.1606</td>
<td>0.1371</td>
<td>0.0000</td>
<td>0.9371</td>
</tr>
<tr>
<td>RDSpendSum Ratio</td>
<td>10,579</td>
<td>0.0485</td>
<td>0.0522</td>
<td>0.0000</td>
<td>0.8856</td>
</tr>
<tr>
<td>Invia</td>
<td>10,579</td>
<td>12.3389</td>
<td>94.6523</td>
<td>0.0000</td>
<td>5,495.00</td>
</tr>
<tr>
<td>IHratio</td>
<td>10,579</td>
<td>0.3514</td>
<td>0.2374</td>
<td>0.0000</td>
<td>0.9866</td>
</tr>
<tr>
<td>Size</td>
<td>10,579</td>
<td>22.1365</td>
<td>1.2645</td>
<td>19.5603</td>
<td>27.4677</td>
</tr>
<tr>
<td>ROE</td>
<td>10,579</td>
<td>0.0689</td>
<td>0.1362</td>
<td>-1.2901</td>
<td>0.4072</td>
</tr>
<tr>
<td>Growth</td>
<td>10,579</td>
<td>0.2106</td>
<td>0.4982</td>
<td>-0.7266</td>
<td>5.8301</td>
</tr>
<tr>
<td>Indep</td>
<td>10,579</td>
<td>0.3770</td>
<td>0.0538</td>
<td>0.3000</td>
<td>0.5714</td>
</tr>
<tr>
<td>TOP10ratio</td>
<td>10,579</td>
<td>0.1524</td>
<td>0.1079</td>
<td>0.0015</td>
<td>0.7942</td>
</tr>
<tr>
<td>Lev</td>
<td>10,579</td>
<td>0.3940</td>
<td>0.1950</td>
<td>0.0268</td>
<td>0.9907</td>
</tr>
<tr>
<td>Dual</td>
<td>10,579</td>
<td>0.3162</td>
<td>0.4650</td>
<td>0.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>TobinQ</td>
<td>10,579</td>
<td>2.2005</td>
<td>1.4627</td>
<td>0.8153</td>
<td>17.7288</td>
</tr>
</tbody>
</table>

Correlation analysis is shown in Table 3.

The absolute value of the correlation coefficient between the main variables is lower than 0.5, indicating that there is no phenomenon of multicollinearity among the main variables selected in this paper, and the final regression results will not produce large deviation. For reasons of space, Table 3 only selects part of the variables to show.

<table>
<thead>
<tr>
<th>Invia</th>
<th>1.000</th>
</tr>
</thead>
<tbody>
<tr>
<td>IHratio</td>
<td>0.061***</td>
</tr>
<tr>
<td>Size</td>
<td>0.173***</td>
</tr>
<tr>
<td>ROE</td>
<td>0.028***</td>
</tr>
<tr>
<td>Growth</td>
<td>-0.008</td>
</tr>
<tr>
<td>Indep</td>
<td>0.005</td>
</tr>
</tbody>
</table>

* , **, and *** indicate statistical significance at the 10%, 5%, and 1% levels. Same as below.
The empirical results are shown in Table 4. The regression coefficients of enterprise innovation input level and institutional shareholding level are -0.021 and -0.062, respectively, and both of them are significant at the level of 1%. In other words, the more institutional shareholding, the less conducive to the improvement of innovation input level, which preliminarily supports the research hypothesis 1a in this paper.

The correlation coefficient between institutional shareholding ratio and the number of high standard patents of listed companies is negative, which is significantly higher than 1%, indicating that institutional investors will inhibit enterprises from carrying out invention patent R&D projects with high technology content but also high risk. However, there is no significant relationship between institutional ownership and low standard innovation. This preliminarily verifies hypothesis 2 in this paper.

<table>
<thead>
<tr>
<th>TABLE IV. REGRESSION RESULTS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Innovation input</strong></td>
</tr>
<tr>
<td>RDSpendSum Ratio</td>
</tr>
<tr>
<td>IHratio</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Size</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>ROE</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Growth</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Indep</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>TOP10 ratio</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Lev</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Dual</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>TobinQ</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Constant</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Observations</td>
</tr>
<tr>
<td>R-squared</td>
</tr>
</tbody>
</table>

*, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels. Same as below.

The robustness test results are shown in Table 5 and 6. In this paper, some control variables are mainly replaced, that is, ROA is used to replace ROE and control year pairs to test the robustness of the dynamic endogeneity model. At the same time, the variables were Winsorized at 1% and 99% quantiles of the sample, and the robust option was added. The results are still robust.
### TABLE V. ROBUSTNESS TEST - ROA REPLACE ROE

<table>
<thead>
<tr>
<th></th>
<th>Invia</th>
<th>RDSpendSumRatio</th>
<th>RDPersonRatio</th>
</tr>
</thead>
<tbody>
<tr>
<td>IHratio</td>
<td>-2.60**</td>
<td>-0.022***</td>
<td>-0.073***</td>
</tr>
<tr>
<td>(-2.31)</td>
<td>(-9.84)</td>
<td>(-10.72)</td>
<td></td>
</tr>
<tr>
<td>Size</td>
<td>6.206***</td>
<td>0.002***</td>
<td>-0.002</td>
</tr>
<tr>
<td>(13.49)</td>
<td>(3.54)</td>
<td>(1.48)</td>
<td></td>
</tr>
<tr>
<td>ROA</td>
<td>24.787***</td>
<td>-0.076***</td>
<td>-0.087***</td>
</tr>
<tr>
<td>(7.64)</td>
<td>(-8.14)</td>
<td>(-3.6)</td>
<td></td>
</tr>
<tr>
<td>Growth</td>
<td>-2.956***</td>
<td>-0.002</td>
<td>0.021***</td>
</tr>
<tr>
<td>(-7.11)</td>
<td>(-1.41)</td>
<td>(5.91)</td>
<td></td>
</tr>
<tr>
<td>Indep</td>
<td>10.986***</td>
<td>0.039***</td>
<td>0.119***</td>
</tr>
<tr>
<td>(-2.29)</td>
<td>(4.88)</td>
<td>(5.01)</td>
<td></td>
</tr>
<tr>
<td>TOP10ratio</td>
<td>-5.684**</td>
<td>-0.034***</td>
<td>-0.109***</td>
</tr>
<tr>
<td>(-2.16)</td>
<td>(-8.95)</td>
<td>(-9.13)</td>
<td></td>
</tr>
<tr>
<td>Lev</td>
<td>-1.178</td>
<td>0.054***</td>
<td>-0.097***</td>
</tr>
<tr>
<td>(-0.91)</td>
<td>(-1.83)</td>
<td>(-10.55)</td>
<td></td>
</tr>
<tr>
<td>Dual</td>
<td>1.023**</td>
<td>0.007***</td>
<td>0.011***</td>
</tr>
<tr>
<td>(2.03)</td>
<td>(7.25)</td>
<td>(3.62)</td>
<td></td>
</tr>
<tr>
<td>TobinQ</td>
<td>1.247***</td>
<td>0.009***</td>
<td>0.021***</td>
</tr>
<tr>
<td>(5.92)</td>
<td>(15.61)</td>
<td>(13.22)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-128.368***</td>
<td>0.111***</td>
<td>0.121***</td>
</tr>
<tr>
<td>(-12.09)</td>
<td>(0.93)</td>
<td>(3.01)</td>
<td></td>
</tr>
<tr>
<td>R-squared</td>
<td>0.078</td>
<td>0.1767</td>
<td>0.1185</td>
</tr>
</tbody>
</table>

*, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels. Same as below.

### TABLE VI. ROBUSTNESS TEST - CONTRAL YEAR

<table>
<thead>
<tr>
<th></th>
<th>Invia</th>
<th>RDSpendSumRatio</th>
<th>RDPersonRatio</th>
</tr>
</thead>
<tbody>
<tr>
<td>IHratio</td>
<td>-2.60**</td>
<td>-0.022***</td>
<td>-0.073***</td>
</tr>
<tr>
<td>(-2.30)</td>
<td>(-9.72)</td>
<td>(-10.74)</td>
<td></td>
</tr>
<tr>
<td>Size</td>
<td>6.161***</td>
<td>0.002***</td>
<td>-0.002</td>
</tr>
<tr>
<td>(13.47)</td>
<td>(3.46)</td>
<td>(1.39)</td>
<td></td>
</tr>
<tr>
<td>ROE</td>
<td>13.588***</td>
<td>-0.044***</td>
<td>-0.044***</td>
</tr>
<tr>
<td>(6.73)</td>
<td>(-8.46)</td>
<td>(-3.24)</td>
<td></td>
</tr>
<tr>
<td>Growth</td>
<td>-3.060***</td>
<td>-0.002</td>
<td>0.020***</td>
</tr>
<tr>
<td>(-7.30)</td>
<td>(-1.49)</td>
<td>(5.75)</td>
<td></td>
</tr>
<tr>
<td>Indep</td>
<td>11.08**</td>
<td>0.039***</td>
<td>0.119***</td>
</tr>
<tr>
<td>(2.30)</td>
<td>(4.90)</td>
<td>(5.02)</td>
<td></td>
</tr>
<tr>
<td>TOP10ratio</td>
<td>-5.606**</td>
<td>-0.035***</td>
<td>-0.111***</td>
</tr>
<tr>
<td>(-2.13)</td>
<td>(-9.31)</td>
<td>(-9.28)</td>
<td></td>
</tr>
<tr>
<td>Lev</td>
<td>-2.376</td>
<td>-0.049***</td>
<td>-0.091***</td>
</tr>
<tr>
<td>(-1.89)</td>
<td>(-17.75)</td>
<td>(-10.45)</td>
<td></td>
</tr>
<tr>
<td>Dual</td>
<td>0.983**</td>
<td>0.007***</td>
<td>0.011***</td>
</tr>
<tr>
<td>(1.95)</td>
<td>(7.35)</td>
<td>(3.63)</td>
<td></td>
</tr>
<tr>
<td>TobinQ</td>
<td>1.255***</td>
<td>0.009***</td>
<td>0.021***</td>
</tr>
<tr>
<td>(5.93)</td>
<td>(15.46)</td>
<td>(13.15)</td>
<td></td>
</tr>
<tr>
<td>Year</td>
<td>control</td>
<td>control</td>
<td>control</td>
</tr>
<tr>
<td>Constant</td>
<td>-126.76***</td>
<td>0.009</td>
<td>0.121***</td>
</tr>
<tr>
<td>(-12.00)</td>
<td>(0.76)</td>
<td>(2.98)</td>
<td></td>
</tr>
</tbody>
</table>
5 CONCLUSION AND DISCUSSION

This paper studies the relationship between institutional investor ownership and firm innovation capability in China's capital market. The main conclusions are as follows: Firstly, institutional investors have a restraining effect on firm innovation output. Secondly, institutional ownership has a negative effect on patent authorization for high standard innovation, but has no significant effect on low standard innovation. In general, the overall institutional shareholding has a certain inhibitory effect on the technological innovation ability of enterprises.

There are still many shortcomings in this paper, so we should reconsider. First, the main dependent variable of this paper, institutional shareholding ratio, is selected as the data of sample enterprises at the end of the year for the convenience of data acquisition. However, institutional shareholding of enterprises is a dynamic process and this dynamic change is not reflected in this paper. Secondly, institutional investors belong to a multivariate aggregate. This paper does not classify institutional investors from heterogeneity. The division criteria are also varied and worthy of further discussion. Finally, the research object of this paper is only listed enterprises, while there are a large number of unlisted smes in China. And different industries are not distinguished, perhaps different industries and enterprise nature will get new findings and conclusions.

REFERENCES

Research on the Competitiveness of China's High-Tech Industry Based on Entropy-TOPSIS Method

Dan Zhao¹, Hanxin Zhang², Qinya Dai², Yu Bai¹*, Lihuang Chen¹

*17802593655@163.com

Advanced Institute of Information Technology, Peking University, Hangzhou 311215, China¹
Hang Zhou Jie Drive Technology CO., LTD, Hangzhou 311223, China²

Abstract: The technology content, product-added value, and labor productivity of high-tech industry are significantly higher than those of traditional industry, so the strong competitiveness of high-tech industry is important for optimizing the industrial structure and improving the quality of economic development. In this paper, we design the high-tech industry competitiveness evaluation system from four dimensions: business environment, R&D investment, innovation achievements, and policy support, and we use the Entropy-TOPSIS method to calculate the high-tech industry competitiveness of Chinese provinces. The conclusion is as follows: (1) The competitiveness of high-tech industry is on an upward trend, and shows a pattern of “the strongest competitiveness in the east, the second in the west, and the weaker in the middle”. (2) The results of high-tech industry competitiveness can be divided into three levels according to the hierarchical clustering method. (3) By comparing the gap of each province in the four dimensions, we find that there are obvious differences in the competitiveness of high-tech industry and unbalanced development in different provinces.

Keywords: High-Tech Industry, Competitiveness Evaluation, Entropy-TOPSIS Method.

1 INTRODUCTION

High-tech industry is an economic entity with knowledge-intensive, capital-intensive, and technology-intensive. Compared with traditional industries, high-tech industry has been greatly improved in terms of production, operation, and technology. As a strategic leading industry to promote the development of national economy, high-tech industry can drive sustained and rapid economic growth and promote industrial upgrading. Therefore, expanding and strengthening high-tech industry is the priority choice for local government to promote economic development. However, due to the differences in economic development level and high-tech industry policies, the regional competitiveness of China's high-tech industry is unbalanced.

At present, the research on the competitiveness of high-tech industry mainly focuses on constructing an evaluation system and calculating it. Chen and Sun (2011) used factor and cluster analysis to evaluate the competitiveness of China's provincial high-tech industry from the perspective of project organization, capital, output, and efficiency [1]. Wang (2014) and Chen (2010) used separately the optimized TOPSIS method and k-means clustering method to calculate the competitiveness of high-tech industry [2][3]. He (2018) compared the
competitiveness of high-tech industry in 31 provinces based on factor analysis and analyzed the competitiveness differences of four economic regions \cite{4}. Deng (2022) used principal component analysis (PCA) to measure the regional competitiveness of Hunan Province from five dimensions: innovation, coordination, green, openness, and sharing \cite{5}.

Based on previous studies, we found those common methods such as factor analysis, clustering method, and PCA are applied to measure the competitiveness of high-tech industry, which is a single measurement method and the results may be inaccurate. To improve the accuracy of the results, we use the comprehensive evaluation method to study the competitiveness of China's high-tech industry, and mainly answer the following questions in this paper: (1) How are the competitiveness of high-tech industry in different provinces? (2) What are the shortcomings of each province in the competitiveness of high-tech industry?

2 THE EVALUATION SYSTEM OF THE COMPETITIVENESS OF HIGH-TECH INDUSTRY

The evaluation system of high-tech industry competitiveness is formed from four dimensions: business environment, R&D investment, innovation achievements, and policy support, and we mainly consider the following aspects:

2.1 Business Environment

The business factors such as a sound legal system, stable macroeconomy, and green ecological environment are the guarantee for the sustainable development of high-tech industry. Therefore, we take the business environment as the most basic dimension to construct the evaluation system of high-tech industry competitiveness.

2.1.1 Marketization Index

The marketization process has significantly promoted the technological progress of China's high-tech industry \cite{6}, and we select the marketization index (x1) to characterize the marketization process which comes from \cite{7}.

2.1.2 External Openness

We measure the degree of openness of each province with the index of external openness (x2), which is constructed from two dimensions of trade openness and investment openness, as shown in Table 1, including export and import dependence \cite{8}\cite{9}, tourism openness \cite{10}, foreign direct investment (FDI), and outward foreign direct investment (OFDI), and calculate openness scores of each province using the entropy method. The data are obtained from the People's Bank of China and the China Statistical Yearbook.

| Table 1. Evaluation index system of China's regional openness. |
|---------------|-----------------|---------------------|
| Dimension     | Indicator       | Indicator description               |
| Trade openness| Export dependency| Export value/regional output value  |
|               | Import dependence| Import value /regional output value |
|               | Tourism openness | International tourism foreign exchange earnings/regional output value |
2.1.3 Digital Financial Inclusion Index

As a representative of the new financial model, the development of digital inclusive finance can reduce the financing cost, and effectively stimulate technological innovation of small and medium-sized enterprises. We use the digital financial inclusion index (x3) to represent the availability of financial services.

2.1.4 Nighttime Light

Nighttime light data record artificial light on the earth's surface and can characterize the intensity of human activities and the urbanization process, which is a good data source for studying human activities. It has been shown that the trend of nighttime lighting data coincides with the trend of GDP to a certain extent and correlates with the urbanization rate, the ratio of three industries, and the number of the population. At present, nighttime lighting data have been widely used in many fields, such as monitoring economic development, urban development, and energy consumption. The nighttime light data (x4) used in this paper comes from the "NPP-VIIRS-like NTL Data" produced by Chen et al. (2021).

2.1.5 Pollutant Emission Index

A good ecological environment is one of the factors that enhance competitiveness and contribute to sustainable development. We use the pollutant emission index (x5) to characterize the environmental quality. The main components of environmental pollutants are industrial wastewater emissions, industrial sulfur dioxide emissions, and industrial smoke emissions, so the pollutant emission index is calculated based on them and the calculation method refers to and . The larger this index indicates more pollution emissions and worse environmental quality, so this index is a negative indicator. This data comes from China Statistical Yearbook and China Environmental Statistical Yearbook.

2.2 R&D Investment

The development of high-tech industry is based on continuous R&D, and there is a correlation between R&D investment and the competitiveness of high-tech industry. The intensity of R&D investment is much higher than other industries, and elements such as talent reserve and research funding can influence the development of high-tech industry. We select R&D personnel full-time equivalent (x6), technology investment acquisition and transformation share (x7), and average R&D project funding (x8) to characterize the R&D investment size of talent, technology, and capital.

2.3 Innovation Achievements

The development of high-tech industry is driven by innovation, and the indicators selected for the innovation achievements dimension in this paper are shown in Table 2. The higher the revenue ratio of high-tech industry (x9) is, the better the development of high-tech industry is in the region. The profit ratio of high-tech industry (x10) reflects the direct economic output of
R&D in high-tech industry. The innovation activities carried out by enterprises are only a means to enhance the competitiveness of the industry, and their ultimate goal is to obtain high profits. The high-tech new product expenditure-to-income ratio ($x_{11}$) reflects the efficiency of new product R&D, which is a negative indicator, that is, the lower this index, the higher the R&D efficiency of new products.

The change in high-tech product exports ($x_{12}$) reflects the international market competitiveness of high-tech industry. High-tech products expand the scale of output at low cost which has become a new growth point for China. R&D patents per capita ($x_{13}$) and trademark applications per enterprise ($x_{14}$) reflect the indirect economic output of R&D in high-tech industry.

### 2.4 Policy Support

Policy support is an important means for the government to stimulate the development of high-tech industry. The government encourages enterprises to carry out innovation activities and increase R&D investment from the aspects of R&D support policy ($x_{15}$) and intellectual property protection ($x_{16}$).

The proportion of government funds in the internal expenditure of R&D funds in high-tech industry reflects the willingness and ability of the government in promoting special expenditure in high-tech industry. Due to the public goods attribute of high-tech products, the optimal allocation of resources in high-tech industry cannot be achieved only by the market mechanism. In the case that individuals or enterprises cannot afford the huge risk brought by investment in basic research, the government directly sponsors the development of basic research through financial allocation, and enterprises make breakthroughs by reinventing products on basic research results. The proportion of technology market turnover to GDP not only reflects the strength of intellectual property protection but also reflects the marketization of high-tech industry.

In summary, we construct the evaluation index system of high-tech industry competitiveness from four dimensions, as shown in Table 2.

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Index</th>
<th>Formula</th>
<th>Var</th>
</tr>
</thead>
<tbody>
<tr>
<td>Business environment</td>
<td>Marketization index</td>
<td>Marketization index of China’s provinces: NERI report 2021</td>
<td>$x_{1(+)}$</td>
</tr>
<tr>
<td></td>
<td>External openness</td>
<td>Calculated by entropy method</td>
<td>$x_{2(+)}$</td>
</tr>
<tr>
<td></td>
<td>Digital financial inclusion index</td>
<td>Digital Finance Research Center of Peking University</td>
<td>$x_{3(+)}$</td>
</tr>
<tr>
<td></td>
<td>Nighttime Light</td>
<td>NPP-VIIRS-like NTL Data</td>
<td>$x_{4(+)}$</td>
</tr>
<tr>
<td>Pollutant emission index</td>
<td>Weights of three pollutants and standardized product</td>
<td>$x_{5(-)}$</td>
<td></td>
</tr>
<tr>
<td>R&amp;D investme</td>
<td>R&amp;D personnel full-time equivalent</td>
<td>The sum of the workload of full-time and part-time personnel</td>
<td>$x_{6(+)}$</td>
</tr>
</tbody>
</table>
3  COMPREHENSIVE EVALUATION MODEL

3.1  Entropy Method

The entropy method determines the weight of the evaluation index. If the weight is higher, the more discrete the data set is and the more information it contains. The calculation steps are as follows.

Step 1: Positive standardization of data.

Establishing a decision matrix $D$ of $n$ evaluation indexes and $m$ evaluation object:

$$D = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix}$$

(1)

To avoid the influence caused by unit differences between different indicators, the data are standardized before the weights are determined, and the normalization matrix $R=[r_{ij}]_{m\times n}$ is calculated. The formulas are as follows:
where $r_{ij}$ is the standardized value in the range $[0,1]$. If $r_{ij}$ is positive, formula (2) is used to standardize the data set. If $r_{ij}$ is negative, formula (3) is used.

Step 2: Calculate the proportion $Y_{ij}$:

$$Y_{ij} = \frac{X_{ij}}{\sum X_{ij}}, 1 \leq i \leq m, 1 \leq j \leq n$$

Step 3: Calculate the entropy $e_j$:

$$e_j = -\frac{1}{\ln m} \sum (Y_{ij} \ast \ln Y_{ij}), 1 \leq i \leq m, 1 \leq j \leq n, 0 \leq e_j \leq 1$$

Step 4: Calculate the variance coefficient $d_j$:

$$d_j = 1 - e_j, 0 \leq d_j \leq 1$$

Step 5: Calculate the objective weight $w_j$:

$$w_j = \frac{d_j}{\sum d_j}, 0 \leq d_j \leq 1, \sum_{j=1}^{n} w_j = 1$$

### 3.2 TOPSIS Method

The TOPSIS method selects the solution whose objective value is with the shortest distance from the positive ideal solution and the farthest distance from the negative ideal solution as the best alternative [24]. The specific steps are as follows:

Step 1: The standardized matrix $R=\{r_{ij}\}_{m \times n}$ and the weight vector $\mathbf{w}=(w_1, w_2, \ldots, w_n)^T$ constitute the weighted decision matrix $Z=\{z_{ij}\}_{m \times n}$ as follows, where $z_{ij}=r_{ij} \ast w_j$, $1 \leq i \leq m$, $1 \leq j \leq n$ and $w_j$ are determined based on the entropy method.

$$Z = \begin{bmatrix}
  \begin{array}{cccc}
  w_1r_{11} & w_2r_{12} & \cdots & w_nr_{1n} \\
  \vdots & \vdots & \ddots & \vdots \\
  w_1r_{m1} & x_{m2} & \cdots & w_nr_{mn}
  \end{array}
\end{bmatrix}_{m \times n}$$
Step 2: Calculate the positive ideal solution $Z^+$ and the negative ideal solution $Z^-$:

$$Z^+ = \{ \max z_{ij} | j \in J \} = (z_1^+, z_2^+, \cdots, z_n^+) \quad (9)$$

$$Z^- = \{ \min z_{ij} | j \in J \} = (z_1^-, z_2^-, \cdots, z_n^-) \quad (10)$$

Step 3: Calculate the distance from each evaluation target to $Z^+$ and $Z^-$ and obtain the maximum distance $S_{i+}$ and the minimum distance $S_{i-}$:

$$S_{i+} = \sqrt{\sum_{j=1}^{n} (z_{ij} - z_j^+)^2}$$

$$S_{i-} = \sqrt{\sum_{j=1}^{n} (z_{ij} - z_j^-)^2} \quad (12)$$

Step 4: Calculate the relative closeness $D_i$:

$$D_i = \frac{S_i^-}{S_i^- + S_i^+} \quad (13)$$

It can be seen that the relative closeness $D_i$ is between 0 and 1 from the formula (13). The greater the $D_i$, the closer the evaluation objective is to the optimal solution.

### 3.3 Data Sources and Processing

The research object of this paper is provinces in China, but Hong Kong, Macao, Taiwan, and Tibet have a large amount of missing data, so they are not included in this study, and the research time range is 2011-2019. The data of the selected indicators in three dimensions of R&D investment, innovation achievements, and policy support are mainly from China Statistical Yearbook (2011-2019), and China High Technology Industry Statistical Yearbook (2011-2019). For the missing data, we used interpolation to fill in the data.

### 4 ESTIMATION RESULTS OF REGIONAL HIGH-TECH INDUSTRY COMPETITIVENESS

#### 4.1 Competitiveness Score of High-Tech Industry in Each Province of China

The competitiveness score of high-tech industry in each province from 2011 to 2019 is measured based on Entropy-TOPSIS method, and the results are shown in Table 3.
Table 3. Competitiveness score of high-tech industry in each province from 2011 to 2019.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Liaoning</td>
<td>0.135</td>
<td>0.14</td>
<td>0.127</td>
<td>0.139</td>
<td>0.125</td>
<td>0.137</td>
<td>0.15</td>
<td>0.158</td>
<td>0.161</td>
<td>0.141</td>
</tr>
<tr>
<td>Beijing</td>
<td>0.411</td>
<td>0.439</td>
<td>0.444</td>
<td>0.45</td>
<td>0.448</td>
<td>0.457</td>
<td>0.474</td>
<td>0.482</td>
<td>0.485</td>
<td>0.454</td>
</tr>
<tr>
<td>Tianjin</td>
<td>0.219</td>
<td>0.224</td>
<td>0.24</td>
<td>0.257</td>
<td>0.244</td>
<td>0.255</td>
<td>0.255</td>
<td>0.254</td>
<td>0.327</td>
<td>0.253</td>
</tr>
<tr>
<td>Hebei</td>
<td>0.073</td>
<td>0.077</td>
<td>0.077</td>
<td>0.08</td>
<td>0.086</td>
<td>0.089</td>
<td>0.109</td>
<td>0.127</td>
<td>0.138</td>
<td>0.095</td>
</tr>
<tr>
<td>Shanghai</td>
<td>0.365</td>
<td>0.33</td>
<td>0.393</td>
<td>0.389</td>
<td>0.408</td>
<td>0.408</td>
<td>0.426</td>
<td>0.44</td>
<td>0.488</td>
<td>0.405</td>
</tr>
<tr>
<td>Jiangsu</td>
<td>0.266</td>
<td>0.273</td>
<td>0.275</td>
<td>0.268</td>
<td>0.271</td>
<td>0.275</td>
<td>0.271</td>
<td>0.279</td>
<td>0.298</td>
<td>0.275</td>
</tr>
<tr>
<td>Zhejiang</td>
<td>0.172</td>
<td>0.183</td>
<td>0.188</td>
<td>0.192</td>
<td>0.195</td>
<td>0.195</td>
<td>0.208</td>
<td>0.231</td>
<td>0.239</td>
<td>0.2</td>
</tr>
<tr>
<td>Fujian</td>
<td>0.163</td>
<td>0.156</td>
<td>0.16</td>
<td>0.155</td>
<td>0.153</td>
<td>0.191</td>
<td>0.193</td>
<td>0.192</td>
<td>0.183</td>
<td>0.172</td>
</tr>
<tr>
<td>Shandong</td>
<td>0.116</td>
<td>0.128</td>
<td>0.141</td>
<td>0.142</td>
<td>0.146</td>
<td>0.149</td>
<td>0.153</td>
<td>0.161</td>
<td>0.173</td>
<td>0.145</td>
</tr>
<tr>
<td>Guangdong</td>
<td>0.344</td>
<td>0.376</td>
<td>0.37</td>
<td>0.366</td>
<td>0.363</td>
<td>0.365</td>
<td>0.37</td>
<td>0.438</td>
<td>0.438</td>
<td>0.381</td>
</tr>
<tr>
<td>Guangxi</td>
<td>0.11</td>
<td>0.086</td>
<td>0.106</td>
<td>0.101</td>
<td>0.109</td>
<td>0.119</td>
<td>0.129</td>
<td>0.14</td>
<td>0.147</td>
<td>0.116</td>
</tr>
<tr>
<td>Hainan</td>
<td>0.201</td>
<td>0.141</td>
<td>0.156</td>
<td>0.154</td>
<td>0.137</td>
<td>0.171</td>
<td>0.224</td>
<td>0.221</td>
<td>0.223</td>
<td>0.181</td>
</tr>
<tr>
<td>Eastern</td>
<td>0.215</td>
<td>0.213</td>
<td>0.223</td>
<td>0.224</td>
<td>0.224</td>
<td>0.234</td>
<td>0.247</td>
<td>0.260</td>
<td>0.275</td>
<td>0.235</td>
</tr>
<tr>
<td>Shanxi</td>
<td>0.058</td>
<td>0.086</td>
<td>0.105</td>
<td>0.122</td>
<td>0.118</td>
<td>0.135</td>
<td>0.139</td>
<td>0.15</td>
<td>0.194</td>
<td>0.123</td>
</tr>
<tr>
<td>Inner Mongolia</td>
<td>0.139</td>
<td>0.056</td>
<td>0.065</td>
<td>0.079</td>
<td>0.09</td>
<td>0.095</td>
<td>0.122</td>
<td>0.154</td>
<td>0.157</td>
<td>0.106</td>
</tr>
<tr>
<td>Jiangxi</td>
<td>0.098</td>
<td>0.143</td>
<td>0.118</td>
<td>0.126</td>
<td>0.142</td>
<td>0.129</td>
<td>0.14</td>
<td>0.156</td>
<td>0.164</td>
<td>0.135</td>
</tr>
<tr>
<td>Jilin</td>
<td>0.137</td>
<td>0.153</td>
<td>0.169</td>
<td>0.123</td>
<td>0.149</td>
<td>0.128</td>
<td>0.144</td>
<td>0.174</td>
<td>0.222</td>
<td>0.155</td>
</tr>
<tr>
<td>Heilongjiang</td>
<td>0.122</td>
<td>0.145</td>
<td>0.152</td>
<td>0.139</td>
<td>0.149</td>
<td>0.143</td>
<td>0.162</td>
<td>0.178</td>
<td>0.192</td>
<td>0.154</td>
</tr>
<tr>
<td>Hubei</td>
<td>0.1</td>
<td>0.106</td>
<td>0.118</td>
<td>0.13</td>
<td>0.144</td>
<td>0.152</td>
<td>0.165</td>
<td>0.171</td>
<td>0.177</td>
<td>0.14</td>
</tr>
<tr>
<td>Hunan</td>
<td>0.14</td>
<td>0.131</td>
<td>0.127</td>
<td>0.12</td>
<td>0.125</td>
<td>0.134</td>
<td>0.135</td>
<td>0.123</td>
<td>0.134</td>
<td>0.13</td>
</tr>
<tr>
<td>Anhui</td>
<td>0.108</td>
<td>0.113</td>
<td>0.122</td>
<td>0.125</td>
<td>0.129</td>
<td>0.134</td>
<td>0.145</td>
<td>0.166</td>
<td>0.17</td>
<td>0.135</td>
</tr>
<tr>
<td>Henan</td>
<td>0.106</td>
<td>0.119</td>
<td>0.13</td>
<td>0.13</td>
<td>0.15</td>
<td>0.155</td>
<td>0.158</td>
<td>0.168</td>
<td>0.195</td>
<td>0.146</td>
</tr>
<tr>
<td>Middle</td>
<td>0.112</td>
<td>0.117</td>
<td>0.123</td>
<td>0.122</td>
<td>0.133</td>
<td>0.134</td>
<td>0.146</td>
<td>0.160</td>
<td>0.178</td>
<td>0.136</td>
</tr>
<tr>
<td>Chongqing</td>
<td>0.118</td>
<td>0.141</td>
<td>0.158</td>
<td>0.165</td>
<td>0.17</td>
<td>0.182</td>
<td>0.187</td>
<td>0.193</td>
<td>0.206</td>
<td>0.169</td>
</tr>
<tr>
<td>Sichuan</td>
<td>0.194</td>
<td>0.208</td>
<td>0.173</td>
<td>0.184</td>
<td>0.192</td>
<td>0.193</td>
<td>0.204</td>
<td>0.222</td>
<td>0.219</td>
<td>0.199</td>
</tr>
<tr>
<td>Guizhou</td>
<td>0.095</td>
<td>0.096</td>
<td>0.1</td>
<td>0.104</td>
<td>0.124</td>
<td>0.105</td>
<td>0.148</td>
<td>0.142</td>
<td>0.192</td>
<td>0.123</td>
</tr>
<tr>
<td>Yunnan</td>
<td>0.117</td>
<td>0.123</td>
<td>0.133</td>
<td>0.127</td>
<td>0.161</td>
<td>0.126</td>
<td>0.141</td>
<td>0.148</td>
<td>0.156</td>
<td>0.137</td>
</tr>
<tr>
<td>Shaanxi</td>
<td>0.191</td>
<td>0.198</td>
<td>0.218</td>
<td>0.232</td>
<td>0.249</td>
<td>0.253</td>
<td>0.256</td>
<td>0.263</td>
<td>0.309</td>
<td>0.241</td>
</tr>
<tr>
<td>Gansu</td>
<td>0.117</td>
<td>0.126</td>
<td>0.122</td>
<td>0.14</td>
<td>0.132</td>
<td>0.136</td>
<td>0.152</td>
<td>0.174</td>
<td>0.194</td>
<td>0.144</td>
</tr>
<tr>
<td>Qinghai</td>
<td>0.081</td>
<td>0.089</td>
<td>0.101</td>
<td>0.099</td>
<td>0.125</td>
<td>0.14</td>
<td>0.148</td>
<td>0.182</td>
<td>0.145</td>
<td>0.123</td>
</tr>
<tr>
<td>Ningxia</td>
<td>0.106</td>
<td>0.115</td>
<td>0.078</td>
<td>0.093</td>
<td>0.121</td>
<td>0.107</td>
<td>0.141</td>
<td>0.129</td>
<td>0.138</td>
<td>0.114</td>
</tr>
<tr>
<td>Xinjiang</td>
<td>0.129</td>
<td>0.152</td>
<td>0.159</td>
<td>0.165</td>
<td>0.158</td>
<td>0.17</td>
<td>0.185</td>
<td>0.198</td>
<td>0.224</td>
<td>0.171</td>
</tr>
<tr>
<td>Western</td>
<td>0.128</td>
<td>0.139</td>
<td>0.138</td>
<td>0.145</td>
<td>0.159</td>
<td>0.157</td>
<td>0.174</td>
<td>0.183</td>
<td>0.198</td>
<td>0.158</td>
</tr>
</tbody>
</table>

We map the high-tech industry competitiveness score of each province with ArcGIS software according to Table 3. As shown in Figure 1(a)-(c), during the study period, the competitiveness of high-tech industry is on an upward trend. We also conclude from the average score of high-tech industry competitiveness, as shown in Figure 1(d), the competitiveness of high-tech industry in the eastern, middle, and western regions has obvious step-like distribution characteristics, showing the pattern of “the strongest competitiveness in the east, the second in the west, and the weaker in the middle”.
Figure 1. Evolution of high-tech industry competitiveness in 2011, 2015, 2019 and average score

4.2 The Hierarchical Clustering Analysis of High-Tech Industry Competitiveness in China's Provinces

The main purpose of hierarchical clustering is to identify variables on some similar or dissimilar features and to divide the variables into several categories according to these features, so that the individuals within the clusters are as similar as possible, while individuals between clusters are as different as possible. We apply SPSS 22.0 analysis tool to cluster the high-tech industry competitiveness score levels of 30 provinces from 2011-2019 based on Euclidean distance and use the Ward method to obtain the distance between different categories. The hierarchical clustering results are shown in Table 4.

Comparing the results of hierarchical clustering analysis with the results obtained by the Entropy-TOPSIS method, it is found that among the three groups of results divided by hierarchical clustering, the scores within each group are similar, and there are no exceptions. Therefore, the clustering results further verify the rationality of using the Entropy-TOPSIS method to evaluate the competitiveness of high-tech industry.
Table 4. Hierarchical clustering results.

<table>
<thead>
<tr>
<th>Level</th>
<th>Province</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strong competitiveness</td>
<td>Beijing, Shanghai, Guangdong</td>
</tr>
<tr>
<td>General competitiveness</td>
<td>Hainan, Fujian, Xinjiang, Chongqing, Sichuan, Zhejiang, Jiangsu, Shaanxi, Tianjin</td>
</tr>
<tr>
<td>Weak competitiveness</td>
<td>Hunan, Qinghai, Ningxia, Guangxi, Guizhou, Shanxi, Inner Mongolia, Hebei, Jilin, Liaoning, Shandong, Heilongjiang, Yunnan, Jiangxi, Gansu, Anhui, Hubei, Henan,</td>
</tr>
</tbody>
</table>

According to Table 1, the clustering results of the competitiveness of the high-tech industry are divided into three levels: strong competitiveness, general competitiveness, and weak competitiveness. Among 30 provinces, the ranking of Beijing, Shanghai, and Guangdong is always in the top three from 2011 to 2019, and they are classified into the strong competitiveness cluster which has an absolute leading edge. And these three provinces are the core city of the Beijing-Tianjin-Hebei urban agglomeration, the Yangtze River Delta urban agglomeration, and the Pearl River Delta urban agglomeration. Tianjin, Shaanxi, Jiangsu, Zhejiang, Sichuan, Chongqing, Xinjiang, Fujian, and Hainan are classified into the middle-level cluster whose competitiveness of the high-tech industry performs generally. In this cluster, Tianjin, Shaanxi, and Jiangsu are slightly better than other regions, but there is still a certain gap compared with the strong competitive cluster. The remaining regions have weaker competitiveness in high-tech industries.

4.3 Dimensional Analysis of High-Tech Industry Competitiveness

We can clearly understand the shortcomings of the province in terms of high-tech industry competitiveness by comparing the gaps in the four dimensions, which will help the province to find and fill gaps. To observe the differences between the four dimensions of high-tech industry competitiveness of each province more intuitively, a radar chart is drawn in this section and shown in Figure 2.
In Figure 2, the closer the point is to the center of the circle, the smaller the corresponding value is. With the increase of the year, each point gradually spreads outward, which means the corresponding value is getting larger. According to Figure 2(a), it can be seen that the competitiveness of high-tech industry in various provinces showed a trend of increasing year by year from 2011 to 2019, and there were four extreme points in spatial distribution: Beijing-Tianjin-Hebei region, Yangtze River Delta region, Guangdong and Shaanxi.

The business environment is the premise to ensure the sustainable and stable development of high-tech industry. As can be seen from Figure 2(b), the development of the business environment in various provinces in China is extremely unbalanced, with a standard deviation of 0.1290. Shanghai, Beijing, and Tianjin rank at the forefront while Gansu, Guizhou, and Qinghai have the worst business environment. In addition, as the business environment is highly valued in China, many provinces have introduced policies to provide enterprises with more efficient and convenient government services and preferential measures which lead to significant growth in the business environment indicator dimension.

From Figure 2(c), it can be seen that the level of R&D investment in Guangdong is significantly higher than that in other provinces, followed by Jiangsu and Zhejiang, while Hainan, Gansu, and Qinghai have the lowest level of R&D investment. The same problem of uneven development among provinces exists in the dimension of R&D investment, with a standard deviation of 0.1521. In addition, Gansu, Heilongjiang, Ningxia, Jiangxi, Guangxi, Yunnan, Jilin, Shanxi, and Shaanxi have achieved large growth in R&D investment in the past few years, and these provinces have relatively high potential in the future development of high-tech industry.

It can be seen from Figure 2(d) that the innovation achievements of Guangdong, Jiangsu, and Shanghai are among the top three, and the innovation achievements of Chongqing, Xinjiang, Beijing, Sichuan, Henan, Shaanxi, Tianjin, Shanxi, Zhejiang are also high, and their standard deviation is 0.0836, the development imbalance of innovation achievements is weaker than other dimensions.
It can be seen from Figure 2(e) that Beijing, Shaanxi, Shanghai, and Sichuan have the highest degree of policy support. The standard deviation of this dimension is 0.1599, and the development of the policy support dimension is the most uneven. Given this, for provinces with weak high-tech industry competitiveness, the local government can refer to the relevant policies of provinces with strong competitiveness to improve the competitiveness of high-tech industry.

5 MAIN CONCLUSIONS AND SUGGESTIONS

In this paper, we design a high-tech industry competitiveness evaluation index system from four dimensions, including business environment, R&D investment, innovation achievements, and policy support. The scores of each province are calculated by Entropy-TOPSIS method. The conclusion is as follows:

(1) The competitiveness of high-tech industry is on an upward trend, and the pattern is presented as “the strongest competitiveness in the east, the second in the west, and the weaker in the middle”.

(2) According to the hierarchical clustering method, the clustering results of high-tech industry competitiveness can be divided into three grades: strong competitiveness, general competitiveness, and weak competitiveness.

(3) By comparing the gap of each province in the four dimensions, we find that there are obvious differences in the competitiveness of high-tech industry and unbalanced development in different provinces.

According to the research conclusions, we put forward the following suggestions: in terms of policy support, the government should encourage enterprises to deduct tax proportionally with the transformation of innovative achievements and reduce the cost of land use for scientific research; in terms of scientific research innovation, the government should set up special financial funds to reward and support related industries; in terms of talent team construction, the government should increase the force of the high-tech talents introduction, and give preference to senior talents in employment subsidies and housing subsidies.

REFERENCES


A Critical Analysis on the Cost Planning in Building Project Success: A Theoretical Review

Zhenquan Zhou1,2,a, Deprizon Syamsunur3, b*, Xinyi Wang1,c

Email: zhenquanzhou@tfswufe.edu.cn, *Corresponding author: deprizon@ucsiuniversity.edu.my, wangxinyi@tfswufe.edu.cn

School of Architecture and Engineering, Tianfu College of SWUFE, Mianyang, 621050, China1
Faculty of Engineering, Technology and Built Environment, UCSI University, Kuala Lumpur, 56000, Malaysia2
Department of Civil Engineering, Faculty of Engineering, Technology and Built Environment, UCSI University, Kuala Lumpur, 56000, Malaysia3

Abstract: The purpose of this article is to discuss how the relationship between project success and cost planning with management techniques. There are three practical ways in cost planning for financial control in this situation: the positive efforts made by cost managers, the cost planning and control process, and practical preventative and corrective actions. The relevant data and information are collected from professional journal publications, lecture notes, reputable web resources, and the reliable experience of construction industry specialists executing cost planning and associated activities. The critical point of this paper is that cost planning can significantly improve the project’s financial performance. For instance, cost planning is likely to reduce project variations and corresponding financial savings. It may also contribute to developing investor confidence by effectively managing total project expenditures. In general, this report analyses the topic comprehensively to understand the importance of cost planning and how this approach positively influences project performance, especially for cost management effectiveness. Furthermore, the positive impact of preventative and corrective activities on the project success is analysed.

Keywords: Cost Planning, Cost Manager, Construction Project, Financial Success, Corrective and Preventative Action.

1 INTRODUCTION

In most situations, a project can be seen as a temporary effort undertaken to generate an exceptional outcome[1]. A group of PricewaterhouseCoopers researchers present that over 97% of organisations globally failed to complete their projects correctly during the previous five years. As a result, an increasing number of managers have attempted to develop a practical approach for increasing the probability of project success in recent years. To a large extent, cost planning has been widely recognised and used by cost managers as an effective technique for
guaranteeing the efficient execution and completion of projects under the estimated plan\[1\]. The next part of this article examines fundamental construction concepts, such as the essential criterion for the building project success and the primary role of the cost manager. Additionally, the third and fourth parts illustrate the common influences of cost planning in building projects and how this approach considerably boosts the possibility of developing a successful building project. Moreover, the final body section explains the critical repercussions of a corrective and preventative way on the project result and success.

2 GENERAL BACKGROUND INFORMATION

2.1 The Project Success

First, it is necessary to understand the definition and significant project success criteria to ensure satisfying results. According to the Project Management Institute, a successful project should accomplish all proposed project goals and assure product quality while adhering to the estimated budget, timeline, and scope limitations\[1\]. Under this definition, the fundamental criteria for project success comply with the project scope, budget, schedule, quality, and safety requirements. For one reason, the triple constraints of time, money, and quality are likely to significantly determine the success of a building project\[3\]. For example, the Titanic construction project could be deemed a failure due to its lengthy delay, higher cost, and inferior product outcomes\[4\]. Besides, the safety requirement is an indispensable component of project success, especially in the construction industry. If a building project results in death or casualty throughout its operation stage, it is considered a severe project failure. For instance, the Third Water Tunnel project in New York City, the most outstanding infrastructure project of the 1970s, is regarded as a failure due to the 24 deaths during its construction\[5\].

2.2 Cost Manager And Quantity Surveyors

Both cost managers and quantity surveyors are essential to ensuring the success of a building project. To some extent, the functions of cost managers and quantity surveyors are so similar that they may be utilised interchangeably in a small-scale construction project. On the other hand, cost managers are often assigned more complicated tasks than quantity surveyors, especially on many large-scale construction projects\[6\]. Consequently, it is critical to understand their roles in building projects. Construction managers often need to provide professional suggestions on building project expenditures to manage and maintain reasonable budgets. It is worth mentioning that they often conduct a range of tasks during a project. For instance, they often assist clients in determining if the expected costs of the intended project are controllable. It probably helps building stakeholders objectively consider multiple alternatives and make the best decisions possible to maximise project benefits\[6\].
3 IMPORTANCE OF COST PLANNING IN BUILDING PROJECTS

The cost plan document is a significant outcome containing a lot of detailed cost planning considerations and measures. It can effectively support the investors and consumers in grasping the practical financial status and related building timetable [6]. Some building professionals agree that cost planning aims to set up the entire cost parameters and figures accurately and realistically during the project planning stage. As the project progresses, the cost manager regularly updates and reports project expenditures and other financial performance in this document. Cost planning, in general, has been actively used by cost managers in recent decades across different continents [7]. The following three paragraphs illustrate the vital significance of cost planning.

3.1 To Ensure Financial Feasibility

Cost planning is crucial to impact and even impact the project financial performance and success because it focuses on the precise analysis and determination of financial feasibility. For example, many excellent building projects in China and other Asian countries would like to undoubtedly present a set of cost strategies at the project beginning phase by recruiting competent cost managers [6]. One of the primary reasons is that the cost planning process will update and present appropriate information to project owners and financial investors regarding the anticipated and practical costs of the targeted project. This function can provide an acceptable budget baseline for critical stakeholders, enabling them to assess the financial feasibility and pursue the projects objectively in their financial conditions [8]. Based on such practical actions, construction investors, especially for many private building projects, can ascertain the profitability of a planned project, preventing and even eliminating the financial risk before the start of a project.

3.2 To Improve Project Profitability

Secondly, the cost planning process is indispensable for guaranteeing and increasing the profitability of building projects. This technique can scientifically predict the probability and contingency of projected expenses occurring during the building lifecycle. Some supporters agree that cost planning can assist in determining the most opportune time that commences a project and secures essential finance at the suitable building stage early. Hence, the planned project is more profitable, resulting in significant tangible benefits for stakeholders due to the project well-organised timeline [8]. On the other hand, if stakeholders are hesitant to use cost-cutting measures, project owners would enter the proposed ventures blindly, sometimes leading to potential bankruptcy [8]. A typical case is related to the building of the Sydney Opera House, which resulted in a considerable financial loss. Due to the functional manager's inexperience with financial analysis and contingency management strategies, the exorbitant cost plan caused a considerable financial loss. Specifically, the project budget was surpassed by around AUS $97 million [5].
3.3 To Update Financial Information Timely

Finally, cost planning may be a powerful instrument for communicating between building specialists and customers. This strategy enables the proper design of construction project financing and timely updates of the project financial data. In practice, the estimated budget will be adjusted throughout the cost planning process in response to the project’s current conditions and expected changes under the scope of work [8]. In other words, the essential building stakeholders have access to current information on the finance and expenditure of the targeted project. The fact is that current information allows many stakeholders in a project to make reasonable contingency management choices. To a certain extent, Building Information Modeling (BIM) is a unique product that originated from the concept of cost planning. It can correctly monitor and inform the present condition of project finance. For instance, this method was utilised to manage the new Royal Adelaide Hospital's construction costs in South Australia. The strategies adopted were generally pragmatic and beneficial for preventing or addressing current project difficulties [6].

4 THE POSITIVE IMPACTS OF COST PLANNING ON THE FINANCIAL SUCCESS OF A BUILDING PROJECT

In many cases, poor cost management is more likely to result in project failures in actual building projects. The Australian Transport Research Forum reports that only 40% of rail projects are finished on schedule and under budget [7]. Fortunately, it is a fact that some cost planning techniques can contribute to the mitigation of financial risks and project success. The following three paragraphs will explain its benefits on financial success from various viewpoints.

4.1 To Decrease The Project Variations For Saving Extra Costs

A successful building development project involves proper market research and the identification of potential dangers prior to initiating the project [9]. That is, meticulous planning can minimise the likelihood of project revisions during the implementation phase. It is worth noting that scope changes during the construction stage will probably cause higher costs, as seen in Figure 1 [6]. Thus, to maintain the project cost within the allotted budget, it is necessary first to assess the market situation and identify potential hazards. Fortunately, this is a significant step involved in the cost planning process. As part of the cost planning control approach, cost managers will acquire the necessary market information for developing a comprehensive risk management strategy for project finance performance. Thus, in real-world building projects, the cost planning technique is often adopted to ensure that scope changes are kept to a minimum during the project lifecycle. It considerably decreases the risk of financial loss throughout the construction and operation phases. As part of the cost planning process, experienced quantity surveyors provide pertinent market information based on the present state of the property market and historical data [10]. Due to the market analysis report, project owners may make educated decisions about land acquisition, development, finance, and contracting with various professionals. It helps developers examine many viability scenarios for a specific site in a development project and choose the most cost-effective one.
4.2 To Enhance the Confidence of Investors

Sufficient finance is essential for a successful endeavour. Cost planning may also instil a sense of trust for project owners and consumers, encouraging them to engage in and financially support the building projects. This situation is because cost planning assists stakeholders in comprehending technical and financial figures. For consumers unfamiliar with construction and cost management, the easily understandable finance report can assist them in swiftly evaluating the financial position and assessing the project sustainability. This sense of involvement will significantly increase their interest and investment confidence [6]. Dr. Huanyu Wu believes that the cost planning control technique may also assist sponsors and investors in gaining a better understanding of a project's financial status. Consequently, they will feel more secure in taking on or sharing project risks, ensuring that the project finance is sufficient to meet anticipated expenditures during its lifecycle.

4.3 To Control the Total Expenditure Effectively

Additionally, a sensible approach to cost planning allows the cost managers to carefully monitor and manage the project budgets within the customer-established cost cap. Meanwhile, a well-organised cost plan effectively handles contingencies [6]. Consequently, it effectively prevents project delays by addressing most potential risks in advance. On this premise, cost planning may be seen as a critical component of guaranteeing financial success since time equals money in most construction cases. Cost planning can also help keep total expenditures within the final budgets, enhancing financial performance. For example, if the planned project would take longer than 12 months to complete, the cost planning control will consider the cost escalation [11]. Without cost escalation, as occurred with the Sydney Opera House, it is far more likely that the project will run over budgets during the post-construction phase, culminating in a financial catastrophe.
5 IMPACTS THROUGH PREVENTATIVE AND CORRECTIVE MEASURES

Realistic preventive and corrective actions throughout the project life cycle may contribute to the project's success in various aspects. It is clear that corrective and preventive action (CAPA) aims to address the underlying and likely causes of nonconformities and other deficient situations, as shown in Figure 2 [12]. For instance, CAPA should minimise potential dangers and address the fundamental causes of present challenges to ensure building quality [13]. In terms of preventative measures, the building project team should typically develop quality management strategies within the first phase of the construction project. This situation is crucial as the quality plan will include requirements for ensuring that work is directed and performed appropriately. In other words, prior to producing defective items, the project team will specify the product demands and monitor the practical operation. Consequently, it effectively avoids product failure because of pointless labour. On the other hand, if the product has already been finished to inadequate quality, the building team will identify the root causes of the problems and provide a plan for fixing them. For example, if the main problem is a lack of qualified workers, they would acquire experienced personnel to replace the inexperienced operators [13].

6 CONCLUSIONS

To summarise, project success is mainly related to its scope, quality, timeline, cost, and safety. It is vital to incorporate cost managers and quantity surveyors early to develop a successful building project since they have sufficient expertise in construction cost control. Cost planning, specifically, is crucial to financial success since it assures financial viability, project profitability, and the project targeted expenditures during its building lifecycle. Meanwhile, cost planning adds to the financial performance by positively influencing tangible or intangible initiatives. Moreover, building project teams need to adopt suitable preventative and reactive approaches, which considerably increases the likelihood of a successful project. If construction professionals can appropriately apply these principles to real-world projects in the future, the construction industry will experience a rise in the number of successful building projects.

Acknowledgements: It is essential to show appreciation for the School of Architecture and Engineering, Tianfu College of SWUFE, as well as the Faculty of Engineering, Technology
and Built Environment, UCSI University based on the support in completing this paper.

REFERENCES


Features of Digital Payment Systems in China and the Prospects for Their Implementation Around the World

Guoliang Sun¹, Vladimir Grigoriev²
1032208828@pfur.ru¹, pluton67@mail.ru²
Department of National Economy, Peoples' Friendship University of Russia (RUDN) Moscow, Russia¹
Senior Lecturer of Institute of World Economy and Business, Peoples' Friendship University of Russia (RUDN) Moscow, Russia²

Abstract: Today in China, the digital payment system Alipay is actively developing - an important trend in China, which serves more than 1 billion users and 80 million sellers around the world[5]. Unfortunately, today a lot of questions arise among such services. Misunderstanding of how they work, what these systems are for, what new they bring to the market, how security and privacy are ensured for consumers of these services. Thus, the relevance of the topic of the article is determined by the active growth of digital contactless payment technology in China.

Keywords: Information Technology, Payment Systems, Mobile Payment Systems, Contactless Payments, Digital Platform.

1 INTRODUCTION

The Chinese banking system is one of the most technologically advanced in the world, which has long been beyond doubt. Digitalization has made mobile apps, online banking, and the combination of financial and non-financial services a part of the daily life of the entire population of China.

In the past few years, we have seen a stable growth in sales via the Internet, an increase in the share of non-cash payments, and an entry into the electronic payments market. The development of non-cash payments should, of course, be stimulated not only by the Central Bank of China, as a regulator, but also by the participants of the banking market by introducing fintech innovations.

Contactless transactions are becoming an important step towards moving away from cash. The article reveals the essence and significance of the Chinese digital payment system Alipay, the features of its use, and the identification of competitive advantages. The article discusses the mechanism of the payment system in China, notes the problems that it may face in the process of implementation and use.
2 ANALYSIS OF THE CHINESE DIGITAL PAYMENT
PLATFORM ALIPAY

Ant Group was formerly known as Ant Financial and Alipay. The company is a subsidiary of the Chinese Alibaba Group. Ant Group is the most valuable FinTech company in the world and the most expensive unicorn company. The group owns Alipay, the largest digital payment platform in China. The total volume of transactions in June 2020 alone reached 118 trillion yuan\(^5\).

Financial services company Ant Group is the largest mobile payment operator in China. The company also deals with insurance, lending, investments and online banking.

Alipay is the main asset of the Ant Group. The application allows you to pay other people, keep money, make investments, buy insurance, get financing, credit cards. Parents can approve their children’s spending in the app. Anyone can pay for their purchase using a QR code, this is the most popular payment method in China.

Alipay is an all-in-one financial application that is used all over China. At the same time, Alipay works in most countries in Asia, Russia, in some European countries and in North America. These are mainly destinations that are very popular with Chinese tourists.

Alipay is a unique platform, unparalleled in the US and Europe. It is not just a payment system like Visa or Mastercard. This system has a very clear financial management structure and high social value among consumers in China and other countries.

Ant Group, which entered the market 15 years ago as an escrow service, is now valued at $ 150 billion. Therefore, it is not surprising that it has become a cause for concern among the leaders of large Western competitors. Many financial companies are impressed and at the same time excited by the rapid development of Ant Group and its payment service Alipay. Moreover, today it is the main payment instrument in China. With the help of the service, Chinese users pay for purchases, services, take instant microloans, open lines of credit, buy insurance, and save for retirement. In total, 730 million Chinese use Alipay every month, for comparison: PayPal has 346 million users\(^6\).

Another example, in 2020, an American company processed $ 764 billion in payments. During this time, Alipay processed $ 18 trillion in mainland China alone.

The Ant Group company plans to raise $ 34.5 billion during the sale of its shares in the near future. If the company does this, it will reach an estimate of $ 313.3 billion. This will allow it to bypass many large giants in terms of capitalization (Figure 1).
Behind this amount is the powerful structure of the Ant Group, its impact on China's financial sector, financial performance and global risks.

The main competitive pressure on Ant is only Tencent Holdings Ltd. (WePay platform). However, Ant is significantly ahead of its competitor.

The real challenge for the company is to maintain an edge in China's complex rules and regulations in obtaining and maintaining proper business permits.

3 ALIPAY DIGITAL PLATFORM STRUCTURE

3.1 Digital Payments

Alipay has a wide range of financial services and super applications compared to its competitors.

The Alipay platform includes digital payments, digital finance, and digital everyday services (Figure 2).
Alipay payment system provides:

- management of bank accounts;
- buying tickets for transport, ordering food, renting a car;
- storage of digital documents;
- shopping on Taobao, Tmall, Alibaba;
- payment of utility bills\(^8\).

AliPay supports almost all types of mobile devices, as well as seamless integration with online assets such as an e-commerce website. Many payment plugins are already supported or under development.

As a result, the company receives revenue from digital payments. It charges sellers and platforms a 0.1% commission for every transaction carried out. The company charges users for transfers to bank accounts and credit card payments.

### 3.2 Smart Loans

On the lending side, Alipay connects about 100 Chinese banks with consumers and start-ups and small businesses looking for short-term loans. Partners take on debt and credit risk, but take interest on loans, and Ant Group takes commission on transactions.

Since 2014, Ant has been offering a range of unsecured revolving loan products to consumers and merchants. Loan products are aimed at small and medium-sized companies, as well as residents of small Chinese cities who do not have the opportunity to purchase a bank card\(^1\).

Ant Group operates a lending payment platform but relies on third-party financial institutions. In 2020, about 98% of loans on the platform were serviced by intermediaries.

Ant Group uses over 100 models to assess the creditworthiness of a borrower. The company assesses the borrower's payment history, his occupation, the amount of money he wants to borrow and other parameters.

Based on the assessment carried out, the borrower is offered personal loan conditions and loan rates. The indicator of delinquencies in payments in the system is 1.5%, when compared with the United States, it is 2.5% for credit cards.

For small and medium-sized businesses, Ant Group offers fairly flexible business development loans with instant repayment on average at 11.6% per annum. These loans are usually repaid in three months.

Ant Group uses the 3-1-0 principle: 3 minutes are allotted for submitting a questionnaire for borrowed funds, in 1 second the payment service decides to issue a loan, all this with 0 interactions with service consumers.

Although the platform has good availability of private finance, users still need a good credit history. For its formation, there is the Huabei tool - quick loans from $ 7 to $ 7000 for online purchases on Taobao, as well as other platforms, and for responsible payers there is the Jiebei tool, which provides loans for amounts from $ 10 thousand\(^7\).
3.3 Investment Instruments

InvestTech's wealth management product gives consumers and businesses access to 6,000 investment products from over 170 funds. Consumers can open an individual investment account right in the Alipay app. Ant uses artificial intelligence technologies to assess risks and offer clients suitable investment options.

Yu'ebao (Residual Wealth) product allows customers to generate income from unused money in their Alipay digital wallet, while the funds can be spent. The service has a low entry threshold - you only need to deposit 1 yuan ($ 0.15). The yield on the offer is 1.7%, which is higher than most interest rates on debit accounts. In June 2020, it was the largest monetary fund in the world market in terms of assets.

3.4 Insurance

The payment platform offers over 2,000 individual insurance products, including health insurance and retirement benefits. Ant Group was the first to provide insurance products to villagers in China.

The payment platform allows insurers to cheaply access millions of customers and process claims online. Ant Group uses natural language processing technologies to validate insurance claims and identify fraudulent claims.

For example, Ant launched Haoyibao, a cancer insurance service. This is the first insurance in China that a customer can renew for their entire life for 89 yuan ($ 13) per year[7].

Let's look at how the Ant payment system works.

In each of the product areas, the company follows three general principles.

1) Data is superpower

Digital payments are important to Ant Group not because of the revenue, but because of the data that can be obtained. Transactions carry, among other things, information about the demographic, socio-economic and civil status of clients and transfer it to other elements of the financial system.

The same goes for business. The data on the number of orders from the seller for six months can be used to predict future cash flows. This information can be used to build a working model of credit history.

2) The first and best customer is Alibaba

The company is entering ever new areas for itself, primarily with its own product, as does Amazon. For example, Ant's first and foremost retail customer is Alibaba, whose needs and data are the driving force behind Ant's next innovation.

First, it's hard to convince legacy financial institutions to focus on technology. Secondly, before offering the platform to other companies, you need to test it well. Third, by opening the door to third-party clients, Ant transforms itself from a financial institution into a platform - and is freed from external financial regulation. Finally, it gives Ant a first-mover advantage in customer relationships.
3) Classic aggregator

Ant Group's products reflect aggregators in all areas, from lending to insurance. These can be direct customer relationships, zero customer service costs and modularity, and vendor availability due to a large user base.

Ant Group quite often used 1 billion users from the Alipay platform to attract banks, mutual funds and insurers. Their significant impact on demand also means they can reduce supplier margins, leading to affordable products with low entry thresholds.

4 PERFORMANCE OF DIGITAL PLATFORM ALIPAY

The company's performance inspires investors: it is a very fast growing profitable business.

![Figure 3. Ant Group Revenues in 2020](image)

The company's revenues are made up of flows: from digital payments, from credit technologies, insurance services, investment technologies (Figure 3).

In 2019, Ant's revenue was $18 billion - 40.6% more than in 2018. The company continued its growth in 2020, with revenues in the first six months of $10.8 billion, which is $7.9 billion more than in the same period a year earlier, and at the end of 2020, revenues were already $19.6 billion (Figure 4).

![Figure 4. Dynamics of Ant Group's income from the operation of the Alipay payment system](image)
During the global economic crisis caused by the pandemic, the company not only did not lose, but also managed to earn money on the payment system with a great result, which suggests that the company's services are profitable (Figure 5).

![Figure 5. Dynamics of Ant Group's profit from the operation of the Alipay payment system][7]

The company's net profit: in 2018, the figure was $ 4.7 billion, in 2019 - $ 8.4 million, and at the end of 2020 - $ 9.8 billion.

## 5 ALIPAY CUSTOMER LOYALTY SURVEY

With the growth in revenue, the number of users also increased. Monthly active users grew from 499 million at the end of 2017 to 712 million by the end of 2020 (Figure 6).

![Figure 6. Growth in the number of users of the payment system (million users)][6]

According to the survey, customer loyalty to the Alipay payment system remains high. 81% of service users prefer to make online purchases through Alipay, and 77% use the service whenever possible.

Alipay is highly rated by most users, but not everyone can trust it. Only 55% of respondents said they trust the payment service, while 80% said the same about banks. More than half (56%) of those surveyed do not want to send cash deposits to their Alipay accounts, however 69% of users believe that Alipay protects their data better (Figure 7).
Thus, we can note that the growth potential of the use of the payment system is confirmed by the fact that today the digital platform with financial services is one of the most demanded by users not only in China, but also in other countries.

The financial service Alipay has a high degree of individualization and simplicity at the same time, which explains its high demand and an important role in increasing financial inclusion. At the same time, being a product of the fintech industry, the payment service often does not fall into the regulatory framework of the Chinese financial market, and, accordingly, can create regulatory and supervisory challenges. In this regard, the success of the development of mobile financial services depends on the interest of financial regulators in stimulating financial innovation and understanding the needs of users.

This challenge leads to a significant expansion of the role of financial regulators from the traditional function of exercising regulation and supervision towards promoting the development of innovative services and the development of competition between banks, telecommunications companies and payment service providers.

In Ant Group's prospectus, “regulatory risk” accounts for 40% of those listed. The requirements of Chinese regulators are constantly changing, affecting one or the other of the group's financial products. The company also notes China's growing geopolitical tensions with the United States and user privacy issues as significant risk factors.

In the future, as the group expands into international markets, these issues will become more pressing and demanding a strong response. But Ant Group has already overcome all obstacles to financial leadership in China.

Ant Group showcases the characteristics and power of China's IT industry. Thus, Western financial innovation becomes negligible compared to Alipay. While the United States and Europe are focused on a new type of bank - trendy, beautiful and affordable, China has spawned a conglomerate that today surpasses traditional financial systems.

6 CONCLUSION

Let's summarize. Currently, the Alipay payment system operates at a cross-border level (USA, Canada, UK, France, Germany, Italy, Russia, Australia, New Zealand, South Africa), and a wide range of services (payments, financial management, insurance, lending, credit scoring) allows us to consider this payment system as a full-fledged ecosystem of financial services.
Today, the digital platform Alipay allows partner companies to look differently at their business and find new directions for development.

A key element of the successful development of innovations on digital platforms is the presence of such a system that creates conditions for the interaction of all market participants, helps to scale and successfully implement new technological solutions.

Acknowledgments: This paper has been supported by the RUDN University Strategic Academic Leadership Program.

REFERENCES


The Use of Information Technology in the Strategy of Internationalization Of Banking

Yi Yao¹, Vladimir Grigoriev²

1032208710@pfur.ru¹, pluton67@mail.ru²

¹Department of National Economy, Peoples' Friendship University of Russia (RUDN) Moscow, Russia
²Senior Lecturer of Institute of World Economy and Business, Peoples' Friendship University of Russia (RUDN) Moscow, Russia

Abstract: The internationalization of banking is impossible without the intensive introduction of the latest achievements of scientific and technological progress in banking. The relevance of the topic is due to the need to optimize the management of information resources in the banking sector. In terms of their importance, automated information systems nowadays occupy one of the leading places in the banking sector. It is impossible to imagine a single modern bank that would not use computer technology. The rapid growth of scientific and technological progress and new information technologies have a significant impact on the overall assessment of the bank's attractiveness. The development of the technological process allows not only to increase the speed of processing documents and conducting cash transactions, but also to expand the clientele. In turn, the development of information technologies can significantly reduce the distance between the producer and consumer of banking services, increase the competition of banks in the international market and contribute to the development of banking services.

Keywords: Information Technology, World Economy, Internationalization Strategies, Banking, Transnational Banks, Financial Globalization.

1 THEORETICAL APPROACHES TO THE CONCEPT OF BANKING INTERNATIONALIZATION

The banking sector is a key actor in the financial industry. Therefore, the internationalization of commercial banks is an important part of financial globalization and the development of the modern world economy. The internationalization of the banking sector is developing on the basis of entrepreneurship. The theory of the internationalization of entrepreneurship was first proposed by the American economist in international relations Raymond Vernon in 1966. He suggested that the internationalization of entrepreneurship is a process from the development of a new product and its sale on the national market, followed by export to international markets[1].

The internationalization of the banking business is a transition to the management of bank property abroad: the creation of branches, the acquisition of a significant share in foreign banks, international mergers, acquisitions and the creation of bank holdings with foreign banks.
The internationalization of the banking business is manifested in the internationalization of banking operations, products and technologies due to the growth of foreign direct investment of banks.

The internationalization of the banking business is manifested in the following:

- international development of traditional banking services, including international settlements, trade and export financing, foreign exchange transactions;
- international activities in the field of non-traditional banking services, including overseas investment, leasing, factoring, forfeiting, securities and insurance.

Thus, in accordance with international rules, banks are engaged directly or indirectly in the provision of international financial services in the international financial market. The internationalization of banking is expressed in the creation of a global distribution system by opening international branches, providing international banking services. Internationalization of banks includes the following elements:

1) Internationalization of management. Banks in the international market must comply with the rules and laws of international banking, international accounting standards, follow the standards for identifying and assessing risks, conducting international audits, personnel management with the requirements of international banking.

2) Internationalization of personnel. The composition and structure of bank employees of the largest transnational banks should be international.

3) Internationalization of capital. Attracting foreign investors, issuing shares and long-term bonds on the international market.

4) Internationalization of the banking network. Creation of international financial institutions, cooperation with foreign companies, shareholding, mergers, acquisitions

5) Customers in the international market. The bank is obliged to have numerous clients in various countries and regions of the world.

1.1 Methods

In the study, the authors used some methods such as analysis and synthesis, induction and deduction, historical and logical, abstraction and concretization.

2 OPERATIONS OF CHINESE BANKS IN INTERNATIONAL MARKETS

After joining the WTO, China provided the country's banking sector with access to the world economy, which made it possible to increase the pace of development of the Chinese financial system and become a leader in international business. With the growth of financial services and the provision of Chinese multinational banks, the integration process in China and abroad is accelerating. The growth of assets, the volume of transnational settlement transactions in RMB and the volume of joint operations in China and abroad have increased significantly in
recent years. In China, there is a clear hierarchy in all areas of life. The banking sector has three distinct levels:

Table I. Levels of the banking sector in China

<table>
<thead>
<tr>
<th>Rank</th>
<th>Types of banks</th>
<th>Tasks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rank 1</td>
<td>The Central Bank of China (People’s Bank of China), political banks (China Development Bank, Agricultural Development Bank of China, and China Exim Bank).</td>
<td>The main task of this level is to distribute and control financial flows, control monetary funds, and issue money.</td>
</tr>
</tbody>
</table>
| Rank 2 | The four banks that support the banking system are called the "Big Four":
1) Industrial and Commercial Bank of China (ICBC);
2) China Construction Bank Corporation (CCB);
3) The Agricultural Bank of China;
4) Bank of China. | These banks are considered the most reliable. The Big Four have the most assets and make up the entire Chinese banking system. They are widely known in many countries around the world, as well as in the Russian Federation. |
| Rank 3 | Private organizations with little income and their own assets. These are small non-profit financial organizations and post offices. | Service for legal entities and individuals |

The activities and development of large Chinese banks are influenced by the following aspects:

- China is confidently moving towards the construction of a trading power, therefore, the development of international business is underway;
- the national currency of China has become more often used in international settlements;
- the active internationalization of Chinese enterprises continues;
- Chinese banks are diversifying their risks and opening up new opportunities to increase profits by entering international markets.

With the strengthening of China’s economy, the competition between national banks is getting fiercer. As a result, the income of Chinese banks in the national market is decreasing. Banks will start looking for additional income from activities in international markets.

Over the past 10 years, more than 15 Chinese banks have successfully consolidated their presence in international markets.

The process of internationalization of Chinese banks takes place through the introduction of innovations, an increase in the level of professionalism in the banking business, and the expansion of the scale of banks’ activities in international markets.

The internationalization of banks has various organizational forms, such as the opening of representative offices of banks, the creation of agencies, subsidiary credit organizations, branches of credit organizations on the territory of other states. Thus, the transnational activities of banks can be carried out in two ways: (1) the opening of new financial institutions and (2) international mergers and acquisitions of financial institutions. The activities of banks
in international markets depend on the possibilities and prospects of international operations, the host country, and internationalization strategies[13].

3 THE CURRENT STATE OF ICBC IN THE INTERNATIONAL MARKET

Industrial and Commercial Bank of China (ICBC) is the largest bank in China. The bank was founded in 1984. Today ICBC is the most expensive banking brand in the world[7].

Mission of the Bank: excellent quality for clients, excellent services for clients, maximum profit for shareholders, real success for employees.

ICBC's vision is to build a state-of-the-art, world-class financial enterprise with global competitiveness, deliver excellence, adhere to a mission, be a customer favorite, be a leader in innovation, security and discretion, and be people-centered.

Values:
- honesty leads to prosperity.
- honesty, humanity, prudence, innovation and excellence.

The bank has 17,000 internal branches and 41 branches around the world.

ICBC pays attention to Internet banking. It enjoys high returns thanks to its extensive client base and nationally regulated bank deposit rates. ICBC and US bank Wells Fargo are competing against each other for the best bank in the world in terms of market capitalization.

Revenue mainly comes from internal subsidiaries. Its subsidiaries include a leasing company, as well as an asset management company and an insurer, which form joint ventures with foreign companies. ICBC strives to become a global titan alongside such financial groups as JPMorgan Chase and Citigroup from the USA.

The Chinese banking conglomerate ICBC, at the end of 2019, took 1st place among the world's 1000 best banks according to The Banker magazine, took 1st place in the Global 2000 rating according to Forbes, topped the list of Global 500 commercial banks in Fortune for the seventh year in a row and for the fourth year in a row ranked 1st among the top 500 banking brands in brand financing[5].

ICBC attaches great importance to the development and implementation of its development strategy in international markets.

ICBC meets all the necessary world-class standards. The Bank's clients are always satisfied with the business policy.

Despite the difficulties caused by the pandemic around the world, ICBC is rapidly moving towards its intended goal to achieve a positive result. The Bank is reforming its structures by managing assets and capital investments.

The main task of ICBC is the reconstruction of the banking system and for this the Bank needs innovations that will change the negative situation caused by the crisis in the global economy.
Therefore, the Bank is betting on a new strategy for restructuring the entire system of the banking business.

To achieve great success in the international market, ICBC must strengthen its position and become competitive. To do this, it is necessary to introduce business innovations and the latest modern technologies.

In November 2007, the Bank opened a subsidiary company in Moscow - ICBC Moscow. ICBC Moscow strives to always put clients first and offers them a wide range of financial services in Russia.

Financial statements are presented in Figure 1. During the pandemic, ICBC raised its stock prices and remained in positive territory (Figure 2). As of April 2021, the stock is up 0.11%. Thus, ICBC Bank is a reliable partner for companies that cooperate with it, as well as individuals.

ICBC provides commercial banking and financial services.

The bank operates in the following business segments: corporate banking, individual banking, treasury operations and others. Despite the financial and economic crisis caused by COVID-19, financial revenues in 2020 increased by 0.4%, profits in 2020 increased by 0.2% (Figure 1)[7].

![Figure 1. Financial performance of ICBC][7]

![Figure 2 - Shares of the Bank during the pandemic in thousand $][7]
The corporate banking segment provides corporate loans, trade finance, depository activities, corporate wealth management services, custodianship and various types of corporate intermediary services for corporations, government agencies and financial institutions.

The personal banking segment offers individual loans, deposits, card business, wealth management services, and various types of personal intermediary services for individual clients.

The treasury segment covers money market transactions, investment securities, foreign exchange transactions and holding positions in financial derivatives on its own account or on behalf of customers. The “Other” segment includes assets, liabilities, income and expenses that cannot be allocated to a segment. ICBC's clients are Chinese, employees of multinational corporations with Chinese origins and students studying in Russia. Currently, ICBC Moscow offers services such as money transfers, currency exchange, and clearing services between countries. ICBC is in its life cycle. On the life cycle of stability, ICBC operates at maximum efficiency. The balance of the company is between growth and profit. ICBC maintains a corporate culture so that employees and stakeholders feel comfortable and spiritually join the company by participating in its life. There is a clear focus and the company clearly fulfills its main function. At ICBC, there is a balance between flexibility and control. In most models of other stages of life, this stage can be classified as a transition from growth to maturity.

ICBC still has enough investment opportunities to make it attractive to investors. ICBC has achieved an optimal life cycle, balancing three important business elements, namely external factors, internal factors and managerial motives.

In order for ICBC to remain stable, it needs to continue to grow by penetrating the international market using an intensive growth strategy. For example, ICBC may step up its market penetration efforts in urban centers in developing countries. Thanks to this external strategic factor, the company can grow through economic growth and the corresponding expansion of the financial services market in countries with high growth rates. Also, ICBC can diversify its business by offering new financial products to increase its revenue.

4 PECULIARITIES OF INFORMATION TECHNOLOGIES AND PROSPECTS OF THEIR APPLICATION IN THE ICBC INTERNATIONALIZATION STRATEGY

ICBC is positioned to be strong and active in terms of global placement. Thanks to many years of efforts, today the Bank has branches all over the world. Its internationalization strategy is a dynamic expansion. ICBC has successfully completed overseas mergers and acquisitions today, which quickly pushed its overseas expansion strategy forward. In the long process of internationalization, ICBC promotes a strategy of expanding external relations that aims at economic development and accelerates the process of internationalization.

Thanks to the globalization of China's economy, ICBC has made great strides in a short time in terms of expansion in international markets, in countries such as Romania, Vietnam, UAE, Pakistan, France, Belgium, Netherlands, Italy and Spain, Russia[8].
At the present stage, the main tasks of ICBC are: development of an internationalization strategy, strengthening financial support for enterprises entering the international market, strengthening the independence of development of foreign banking institutions.

ICBC’s internationalization strategy in the Russian market is focused on investing in brand awareness activities. Increasing ICBC brand awareness is important to help the Bank enter new consumer segments and increase awareness. ICBC should also introduce consumers to new markets for its banking products. By increasing sales through training new consumer segments in existing and new markets, the Bank leads the Bank to overall business development and growth.

ICBC offers a wide range of financial products:
- provides financial services to Chinese and Russian export and import enterprises;
- invests in projects and financing of enterprises in the real sectors of the economy;
- conducts international payments;
- carries out conversion operations;
- helps Russian enterprises to develop their business in China;
- contributes to the development of trade and economic relations, the organization of investment and business cooperation between Russia and China.

The use of modern information technologies radically affects the internationalization of the Bank and changes business processes, bringing it to a fundamentally new level. Banking technologies are inextricably linked with information technologies, which provide complex business automation.

At the heart of ICBC’s changing IT landscape is the technology platform, which acts as a business environment that enables ecosystem members to interact and create value. The technological platform includes infrastructure, data and tools for their processing and analysis, applications, development tools, API.

ICBC actively applies the latest technologies to create innovative services and services. Methods of working with Big Data were used in the development of a universal chat platform for instant messengers. A retinal-based biometric identification system for self-service devices is being introduced. A system for the development of a mathematical model of cash and cash management in ICBC self-service devices has been implemented, which is designed to reduce device downtime and save on optimizing the stored amounts of cash.

On the basis of an open interface (API), a prototype of a portal has been developed that allows external ICBC partners to manage the life cycle (API) for connecting ICBC-Messenger partners. It also publishes corporate API services, which helps in attracting new partners.

ICBC has introduced the technology of "direct settlement", which made it possible to make payments in the ICBC settlement system in real time.

ICBC has built a basic ICBC product portfolio based on aggregated data. The first deals were made with external clients for products to build a model of propensity to buy and segment the customer base for targeted campaigns.
The Open Data portal was launched - a unique information product based on Big Data technologies, which presents aggregated data on the economic activity of the population and business.

ICBC has established innovation laboratories, which have an average development time of an innovative product of about five months. Among the innovative initiatives that ICBC is testing or implementing are the following:

- accounting and management of powers of attorney, electronic mortgages for real estate, money transfers, accounting of factoring transactions based on blockchain technology;
- building an electronic document management system based on blockchain technology;
- automated construction of work schedules for employees in customer service offices;
- using the technology of automatic communication with clients in text service channels (messengers);
- using smart digital assistants to solve customer problems and a universal platform for chatbots;
- Creation of a new generation mobile banking application based on messenger technology;
- creation of an ecosystem of customer communication, simple communication channels, management of accounts and funds, transfers; research of marketplaces for partners, provision of additional services to clients.

ICBC adopts the experience of advanced world developments in the field of information technology, and also develops its own Programs for the development of the banking sector and improving the quality of service while minimizing costs and risks. Key ICBC programs are shown in Figure 3.

![Figure 3. ICBC technology platform](image)

The main task of the technological platform is to reduce the development time of banking products (time-to-market) and increase the level of automatic processing of transactions.
The new modeling platform will reduce operational risk and allow you to quickly implement models of any complexity into business processes. The implemented project will ensure the continuity of the development and implementation of machine learning models with their subsequent integration into ICBC business processes. The deployed architecture automates the management of models, working with them is carried out in a separate secure test environment, and the changes are automatically transferred to the industrial environment. ICBC Bank managed to build processes for the development of modeling teams, providing them with advanced technologies. For the development of machine learning models, access to corporate data warehouses has been implemented, ready-made images of workplaces for teams with a different stack of machine learning libraries have been configured. Automated processes of transferring models from the development circuit to the application circuit will allow in the future to reduce the time-to-market tasks of introducing machine learning into the decision-making processes of the bank, as well as to reduce the operational risk of manually transferring the logic of models between systems.

5 CONCLUSION

ICBC in the international market can strategically benefit from the development of its international payment clearing services, expand its client base to a global scale.

Since the start of the coronavirus pandemic, mobile capabilities have become a more significant factor when choosing a bank among respondents. ICBC needs to understand which mobile banking features are most valued by consumers and where the Bank is compared to its competitors. This is necessary in order to update digital platforms and stay ahead of the competition.

The strategy of internationalization of the banking business, its management and performance assessment necessitate the modernization of information technology systems in the bank, improve the quality of client work, determine the development strategy of the bank and strategic planning.

The transition to international reporting standards, the reduction of banking risks also predetermine the conditions for choosing directions for the development of the Bank’s automation, choosing a specific information system and planning investments in it.

Acknowledgments: This paper has been supported by the RUDN University Strategic Academic Leadership Program.

REFERENCES


[12] Internationalization of Banks European Cross-border Deals // https://link.springer.com/content/pdf/
[13] Internationalization theories // https://nspolteksby.ac.id/
An Electronic Channel Operation Method for Power Marketing Based on User Behavior Analysis

Qirui Chen*, Wei Zhang, Jianing Xu, Jiali Yu, Ying Jiang

*lingling10594184@163.com

Marketing Service Center of State Grid Zhejiang Electric Power Co., Ltd., Hangzhou, Zhejiang, 310014, China

Abstract: With the abolition of the national industrial and commercial catalogue electricity price, the competition in the electricity sales market has become increasingly fierce, and the importance of electricity marketing channels has become increasingly prominent. In order to solve the problem of single power marketing channels and low adaptability to power customer needs, a power marketing electronic channel operation method based on user behavior analysis is proposed. Firstly, based on the electricity consumption, files and other information of power users, user behavior analysis is carried out to obtain typical user characteristic behaviors. Secondly, based on Internet+, the potential index of electricity marketing electronic channels of power users is analyzed. On this basis, the operation model of electric power marketing electronic channel is established, and the operation cost of electricity is reduced and customer satisfaction is improved by formulating targeted operation strategies. Finally, the operation results of the proposed method in a city show that the operating cost of the method is reduced by 7.6% compared with the traditional method. The application results verify the effectiveness of the proposed method.

Keywords: User behavior analysis, Electricity marketing, Electronic channels, Operational methods, Catalogue electricity prices, Electricity sales market.

1 INTRODUCTION

With the deepening of China's new round of power system reform, the importance of power marketing is increasing [1]. In 2021, after China's National Development and Reform Commission issued a document abolishing the catalogue electricity price for industry and commerce, competition in the industrial and commercial electricity sales market will become increasingly intense [2]. How to improve the level of electricity marketing services and increase the ability of industrial and commercial users to acquire customers is a key element of the research of power supply companies and electricity sales companies [3]. In the traditional power marketing of power supply companies, the combination of offline business halls and online APP is used to provide marketing services for industrial and commercial users, but the operation mode is single, which cannot meet the growing demand for power supply services of industrial and commercial users. Therefore, power supply companies urgently need to carry
out electronic channel expansion services to improve the active service level of power supply companies [4].

In order to solve the problem of insufficient user incentive in power marketing channels, a power marketing electronic channel operation method based on user behavior analysis is proposed [5]. This method collects the non-adjustable load, adjustable load, transferable load and interruptible load of users, and adopts the multi-label classification method to establish the characteristics of typical electricity customers, which improves the granularity of user power consumption behavior analysis [6]. On this basis, the potential index analysis of the electricity customers targeted by this method is carried out, and different types of power services are provided for different types of electricity customers, so as to improve the satisfaction of electricity customers and enhance the stickiness of electricity customers.

2 ELECTRIC POWER MARKETING ELECTRONIC CHANNEL OPERATION MODEL

2.1 Establish the behavior characteristics of electricity customers

The purpose of establishing the behavior characteristics of electricity customers is to portray the characteristics of electricity customers from the aspects of the industry characteristics of electricity customers, the contribution value of electricity customers, the power demand of electricity customers, and the growth characteristics of electricity customers, so as to establish a portrait model of electricity customers in order to formulate accurate power marketing electronic channel operation strategies.

The data set used to establish the characteristics of electricity customers is shown in Table 1

<table>
<thead>
<tr>
<th>serial number</th>
<th>Collect data</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Electricity customer files</td>
</tr>
<tr>
<td>2</td>
<td>96 points daily load curve</td>
</tr>
<tr>
<td>3</td>
<td>Factor</td>
</tr>
<tr>
<td>4</td>
<td>Electricity price information</td>
</tr>
<tr>
<td>5</td>
<td>User default records</td>
</tr>
<tr>
<td>6</td>
<td>Marketing inquiries</td>
</tr>
</tbody>
</table>

The multi-label classification method is a method of electricity customer portrait, which can extend the existing method and can process multi-label data [7], which is suitable for a large number of electricity customer data classification. Therefore, the multi-label classification method is used to classify Table 1.

In multi-label classification, the sample set of electricity customers $G_a$ is:

$$G_a = \sum_{j=1}^{n_i} (U_j + O_j)$$

(1)
Formula: the $n_b$ number of behavioral characteristics of electricity customers; $U_i$ Electricity consumption behavior for different users; $O_i$ For the needs of different users.

User mappings are categorized $K_b$ as:

$$K_b = \sum_{i=1}^{n_b} (W_a G_w \sigma)$$  \hspace{1cm} (2)

Formula: $n_b$ map the number of classifications for electricity customers; $W_a$ Electricity customer labeling behavior for different mapping classifications; $G_w$. A sample of electricity customers who classify different mappings. $\sigma$ feature mapping function for electricity customers.

The feature mapping function of the electricity customer satisfies the following constraints:

$$\sigma \in \varphi$$  \hspace{1cm} (3)

Formula: $\varphi$ Classify boundaries for electricity customers.

According to the start and stop time, use time, power size, and power characteristics of the customer's electrical equipment, identify the user's electricity consumption behavior characteristics. Electricity customers usually include four categories: non-adjustable load, adjustable load, transferable load and interruptible load, and the behavior of electricity customers using electrical appliances is set by formula (2).

1. The load cannot be adjusted, and the time and power of the electricity customer using such electrical appliances cannot be adjusted, such as televisions, electric lights, etc.
2. The load can be adjusted, and the electricity customers cannot stop using electrical appliances, but the power can be adjusted, such as: water heaters, air conditioners, etc.
3. Transferable load. Electricity customers' appliances can be stopped and adjusted to another time, such as office shredders, etc.
4. The load can be interrupted. Loads that can be started and stopped at any time by electrical appliances of consumers, such as electric vehicles.

Taking agricultural and animal husbandry users as an example, according to the daily electricity consumption curves of four load types of electricity customers, a finite propagation (AP) was adopted. Establish typical behavior characteristics of electricity customers, limited to space, the AP clustering process will not be repeated in the paper, and the AP clustering method is detailed in the reference. The power consumption curves of various types of load consumers are shown in Figure 1.
In Figure 1, the non-adjustable load of the agricultural and animal husbandry user is the aquatic oxygenation system, and the power consumed during 24 hours remains unchanged; The adjustable load is used in the morning, afternoon and evening time slots, and the power can be adjusted; The transferable load is the garbage removal robot, which is charged during the noon period; The interruptible load is an electric vehicle, which is charged during the morning time slot.

Establish the behavior characteristics of electricity customers and transmit various load curve information of electricity customers to the potential index analysis of electric power marketing electronic channels to carry out potential index analysis.

2.2 Analysis of the potential index of electronic channels for power marketing

The purpose of the analysis of the potential index of electronic channels of power marketing is to evaluate the potential index of electronic channels such as Pocket Power, State Grid Mall, Alipay, WeChat, 95598 and so on, so as to determine the promotion coefficient of that channel.

Internet + is the latest trend of the current Internet development, the power industry widely uses Internet + technology, and in terms of marketing channels, vigorously promote Internet + marketing, so as to reduce the processing pressure of power supply business halls, reduce the operating costs of power supply companies, realize the digital transformation of power supply companies, and improve the satisfaction of electricity customers.

Firstly, the power marketing electronic channel is divided into $X$ two coordinates, $Y$ where the coordinates represent the acceptability of the power marketing electronic channel to electricity customers, the larger the $X$ acceptability value, the higher the ability to be accepted, and the coordinates indicate the degree of attraction of the power marketing $Y$ electronic channel to electricity customers, the greater the degree of attraction. Explain that the more willing the user is to pay.

The analysis of the potential index of electronic channels for electricity marketing $C_a$ is:
Formula: $n_{\alpha}$ is the acceptability of electric power marketing electronic channels, and $\sigma$ the potential analysis function of power marketing electronic channel potential index; $B_i$ For different power marketing channels.

The electronic marketing channels $B$ are:

$$B = h_a \times h_c$$

Formula: $h_a$ is the acceptability of electricity marketing channels; Degree of attraction $h_c$ for electricity marketing electronic channels.

The electronic channel potential index of power marketing transmits the information of the user's electronic channel potential index to the establishment of the electronic channel operation model.

2.3 Establish an electronic channel operation model

The purpose of establishing the electronic channel operation model is to better formulate the marketing operation plan for electricity customers. First, the electronic channel operation analysis is carried out through data such as region, season, and electricity customer payment information, and then the design of power marketing power operation activities is carried out, and then the operation strategy of the electronic channel of power marketing is formulated.

Innovative diffusion analysis is a method for electronic channel operations, which obtains operational analysis results by simulating the perception of the electricity consumption process by simulating individual electricity customers.

The results of the operational analysis $z_a$ are:

$$Z_a = \frac{1}{[D_a + S_a \times e_a]}$$

Formula: $D_a$ the perceived information of electricity customers; $S_a$ Estimated parameters for electricity customers; $e_a$ is a linear parameter.

Then formulate the operation strategy of power marketing electronic channels. According to the user scope and characteristics covered by the electronic channel of power marketing, realize the promotion of the power grid company’s own channels; The promotion methods include: advertising, posters, self-media publicity, etc., to promote the activity of existing electricity customers and promote the participation of new users.
Establish an electronic channel operation module to transmit electronic channel operation strategy and promotion method information to the operation evaluation of electric power marketing electronic channels.

2.3 Operation evaluation of electronic channels for power marketing

Through the Internet plus power marketing electronic channel operation service, the electronic channels of Pocket Power, WeChat, Alipay, State Grid Mall, and 95598 have been launched, and users use electronic channels to complete electricity consumption business such as payment, electricity information inquiry, and business expansion information handling. Therefore, the coverage of electronic channels for power marketing \(^8\) will be increased, and the power supply service level of power supply companies and the satisfaction of electricity customers will be improved.

Linear regression is a mathematical statistical method used to evaluate the operational effectiveness of electronic channels in power marketing, which can determine the interdependence between multiple variables.

The evaluation relationship \( P_i \) is:

\[
P_i = \phi + \eta l_i + \tau
\]  

(7)

Formula: \( l_i \) Evaluation of independent variables for electricity customers; \( \phi \), \( \eta \) is a regression constant; \( \tau \) is a random perturbation quantity.

3 EXAMPLE ANALYSIS

In order to verify the operation method of electric power marketing electronic channel based on user behavior analysis proposed in this paper, the method proposed in this paper is applied in a municipal power supply company. The number of electricity customers is 1.2 million, residential and agricultural users are 920,000, industrial and commercial users are 275,000, and large industrial users are 5,000. The number of servers in the model mentioned in the article is 10, running with F5 connection, running the operating system as win server, and the processor is Xeon 32-core 3.6G. The running memory is 128G.

3.1 Time-consuming analysis of electronic channel operation plan

The purpose of the time-consuming analysis of the electronic channel operation solution is to verify the operation performance of the proposed model. The indicator is calculated from the time the basic data is entered to the time when the operation plan of the electronic channel for power marketing is generated. The shorter the time of this indicator value, the better the performance of the electronic channel operation plan.

Selecting different user number test sets can verify the time-consuming electronic channel operation scheme of the model under different user numbers, so as to reduce the error caused by the use of a single test set. Select the number of electricity customers as 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 10000, comparing the power marketing electronic channel
operation method based on user behavior analysis and the dimension decomposition method widely used in the industry, the comparison results are shown in the table 2 is shown.

<table>
<thead>
<tr>
<th>Number of electricity customer samples</th>
<th>Model running time(s).</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>User behavior analysis</td>
<td>Dimension decomposition</td>
</tr>
<tr>
<td>1000</td>
<td>1.19</td>
<td>6.26</td>
</tr>
<tr>
<td>2000</td>
<td>2.40</td>
<td>12.55</td>
</tr>
<tr>
<td>3000</td>
<td>3.64</td>
<td>18.85</td>
</tr>
<tr>
<td>4000</td>
<td>4.78</td>
<td>25.05</td>
</tr>
<tr>
<td>5000</td>
<td>5.96</td>
<td>31.51</td>
</tr>
<tr>
<td>6000</td>
<td>7.26</td>
<td>37.69</td>
</tr>
<tr>
<td>7000</td>
<td>8.35</td>
<td>43.84</td>
</tr>
<tr>
<td>8000</td>
<td>9.68</td>
<td>50.20</td>
</tr>
<tr>
<td>10000</td>
<td>11.87</td>
<td>62.59</td>
</tr>
</tbody>
</table>

It can be seen from Table 2 that under the number of different samples of electricity customers, user behavior analysis runs fast in the four parts of "establishing electricity user behavior characteristics", "power marketing electronic channel potential index analysis", "establishing electronic channel operation model" and "power marketing electronic channel operation evaluation", therefore The electronic channel operation method of electric power marketing based on user behavior analysis proposed in this paper takes less time than that of dimensional decomposition method, which shows that the proposed method has better performance in generating electronic channel operation scheme.

3.2 Efficiency analysis of electronic channel operation scheme

The efficiency analysis of electronic channel operation scheme is the core index of electricity customer analysis in the paper. In order to simplify the calculation, the satisfaction rate of electricity customers is used to replace the efficiency analysis of electronic channel operation scheme. If the electricity customers are satisfied with the electronic channel operation scheme, the electronic channel operation scheme is effective. The calculation method of the satisfaction rate of electricity customers is as follows: after the power supply company provides services to electricity customers according to the electronic channel operation plan, they collect whether the customer is satisfied online. If the customer is satisfied, the electronic channel operation scheme is effective, and the ratio between the two is the efficiency of the electronic channel operation scheme. When 10,000 users are selected in 5 regions of a city to compare the growth of electricity consumption in the same period, if the monthly electricity consumption of users is larger than the average level of the previous year, then the users are satisfied. The value range of this index is 0-100%. The higher the index value, the more satisfied the electricity customers are.

Compared with the customer satisfaction of the electronic channel operation method of power marketing based on user behavior analysis and the dimensional decomposition method widely used in the industry, the results are shown in Figure 3.
As can be seen from Figure 2, in different regions of the city, the average satisfaction rate of the electronic channel operation method based on user behavior analysis is 98.3%, which is higher than the average 89.6% of the dimensional decomposition method. It can be seen that the proposed method in this paper has better effect.

3.3 Analysis table of electronic channel operating expenses

The purpose of the operating cost analysis is to measure the direct economic value brought by the method mentioned in the paper. The lower the cost, the better the economic benefit of the operating method.

Select two areas of a city and compare the operating expenses of the proposed method with the traditional method for one month. The comparison results are shown in Table 3:

<table>
<thead>
<tr>
<th>Region</th>
<th>Cost (ten thousand yuan)</th>
<th>Amplitude of fluctuation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Analysis of user behavior</td>
<td>conventional method</td>
</tr>
<tr>
<td>region 1</td>
<td>23.62</td>
<td>25.35</td>
</tr>
<tr>
<td>region 2</td>
<td>24.73</td>
<td>26.67</td>
</tr>
<tr>
<td>subtotal</td>
<td>48.35</td>
<td>52.02</td>
</tr>
</tbody>
</table>

As can be seen from Table 3, after the operation of the electronic channel operation method of power marketing based on user behavior analysis mentioned in this paper, the operation effect of we-media publicity and advertising is good, and the operating cost supported by it decreases by 7.6% compared with the traditional methods.

4 CONCLUSIONS

In order to solve the problem of single power marketing channels and low adaptability to power customer needs, a power marketing electronic channel operation method based on user behavior analysis is proposed. On the basis of analyzing the behavior of electricity customers, the potential index of electric power marketing electronic channels is analyzed, and the
operation model of power marketing electronic channels is established accordingly, and targeted operation strategies are generated to reduce power operating costs and improve electricity customer satisfaction. Finally, the operation of a municipal power supply company verifies the effectiveness of the proposed method.

In the next step, the electronic channel operation method mentioned in the paper will be further studied in combination with the provincial electricity customer data.

REFERENCES

Research on Scientific Management of Living Materials During the Coronavirus Disease 2019 Epidemic

Feng Chen*

e-mail: 1539668727@qq.com

School of Management, Shanghai University of Engineering Science, Shanghai, China

Abstract: The sudden outbreak of the novel coronavirus disease 2019 (COVID-19) at the end of 2019 has had a significant impact on the lives of Chinese people and social and economic development. During the period of epidemic closure, the scientific management of various groups of people has become very important, and the scientific management of living materials is particularly important. This paper uses the BP neural network model to predict and analyze the development trend of the distribution of vegetable bags and the number of people infected with the new crown and finds that the inflection point of the epidemic after the distribution of vegetable bags appears early and shows a clear downward trend. This paper concludes that the scientific management and distribution of living materials during the epidemic can reduce the intensive contact between people, reduce the risk, and achieve the effect of inhibiting the development of the epidemic. This paper concludes the predictive analysis of the application of technical means, which will provide solutions for the emergency response to major public health events in the future.

Keywords: Novel Coronavirus Pneumonia, BP Neural Network, Living Materials.

1 INTRODUCTION

Since 2022, our country has repeatedly experienced large-scale repeated epidemics. In response to large-scale outbreaks, our country has adopted a closed management method to achieve rapid elimination of the epidemic. The new type of coronavirus pneumonia has the characteristics of rapid onset and strong contagion. To reduce the number of infected cases due to the increase in people gathering, most areas in our country have adopted large-scale closed management to block the spread of the virus [1]. During the lockdown period of the epidemic, the scientific management of living materials has become particularly important. To block the secondary spread of the epidemic, we must propose a scientific and easy-to-implement method for the distribution of living materials.

Since March 2022, the situation of the new crown epidemic in Changchun has become more and more serious and gradually spread to the whole city. To effectively control the spread of the
epidemic and protect the lives, health, and safety of the people in Changchun City, the Changchun Municipal Government and units at all levels have actively carried out anti-epidemic measures and taken some effective measures to reduce the spread of the epidemic. To ensure the basic living supply of the masses, Changchun City has formulated a series of measures to ensure the reasonable, timely, and effective distribution of materials. The method of material distribution reduces the intensive contact of people to a certain extent, reduces the transmission route of the new crown virus, and inhibits the development of the new crown epidemic to a certain extent. During the period of material supply, the distribution method, time interval, and transportation of vegetable packs are very important. The shelf life of vegetables is short, the demand is large, and storage is difficult. The government and relevant units need to formulate an effective supply process to reduce losses and enable the delivery of materials, maximize utility and overcome the epidemic situation better and faster.

2 MODEL ESTABLISHMENT AND SOLUTION

2.1 Person Correlation Analysis

1) Introduction to Person Coefficient

The person correlation coefficient is a statistical method proposed by British statistician Pearson in the 20th century to measure the closeness of the correlation between two variables. It analyzes the correlation by measuring the linear relationship between two variables X and Y. The calculation formula is represented by r, and the value range is between -1 and 1. The calculation formula is (1):

\[ r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2\sum_{i=1}^{n}(y_i - \bar{y})^2}} \]  
(1)

where \( x_i \) and \( y_i \) are the sample means.

The meaning of the formula is: when \( r > 0 \), the two variables are positively correlated, for example, the greater the value of the x variable, the greater the value of the y variable; when \( r < 0 \), the two variables are negatively correlated, the larger the value of the x variable, the smaller the value of the y variable will be. That is to say, the larger the absolute value of the correlation coefficient r, the stronger the correlation between the two variables, that is, the closer r is to 1 or -1, the stronger the correlation, and the closer r is to 0, the stronger the correlation is. The weaker. When using person correlation to analyze the relationship between variables, SPSS software is generally used for operation, and the correlation coefficient of r is marked in the obtained result table, and the correlation between variables can be seen; in addition, there is a P value, it is used to indicate the significance level of this analysis. When the p-value is greater than 0.05, there is no need to examine the correlation between variables. When the p-value is less than 0.05, it indicates that there is a correlation between the variables. The r value considers whether the variables are positively or negatively correlated.
2) Analysis of Person correlation results

Using SPSS24 for Person correlation analysis, the variables analyzed in this paper are mainly divided into two categories, one is the number of new coronavirus infections in each district of Changchun City, and the other is the number of vegetable packs distributed in Changchun City. The number of new coronavirus infections in each district was matched with the number of vegetable packs distributed in the district, and 9 correlation analyses were performed. The final results are shown in Table 1.

<table>
<thead>
<tr>
<th>District</th>
<th>Number of vegetable packs issued</th>
<th>Pearson correlation</th>
<th>Significance test (two-tailed)</th>
<th>Number of cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Changchun New District</td>
<td></td>
<td>-0.537*</td>
<td>0.032</td>
<td>16</td>
</tr>
<tr>
<td>Economic Development Zone</td>
<td></td>
<td>-0.472*</td>
<td>0.015</td>
<td>16</td>
</tr>
<tr>
<td>Er dao District</td>
<td></td>
<td>-0.536*</td>
<td>0.033</td>
<td>16</td>
</tr>
<tr>
<td>Green District</td>
<td></td>
<td>-0.275**</td>
<td>0.002</td>
<td>16</td>
</tr>
<tr>
<td>Kuan cheng District</td>
<td></td>
<td>-0.279*</td>
<td>0.026</td>
<td>16</td>
</tr>
<tr>
<td>Jing yue District</td>
<td></td>
<td>-0.582**</td>
<td>0.008</td>
<td>16</td>
</tr>
<tr>
<td>Chaoyang District</td>
<td></td>
<td>-0.542**</td>
<td>0.03</td>
<td>16</td>
</tr>
<tr>
<td>Nan guan District</td>
<td></td>
<td>-0.399**</td>
<td>0.006</td>
<td>16</td>
</tr>
<tr>
<td>Auto open area</td>
<td></td>
<td>0.194</td>
<td>0.472</td>
<td>16</td>
</tr>
</tbody>
</table>

* At the 0.05 level (two-tailed), the correlation is significant
2.2 BP Neural Network Model

BP neural network is a multi-layer forward network model that is widely used and effectively based on an error back-propagation algorithm. It was first proposed in 1986 by a team led by Rumelhart and McClelland. BP neural network can simulate the information transmission mode of human brain neurons, perform nonlinear transformation and regression processing on complex information variables, and obtain highly fitting operation results. By simulating the information transmission between neurons in the brain, after inputting the input variables into the input layer, a linear combination of the input variables will be obtained according to the initial set weights\(^8\). When the weights are continuously modified so that the linear combination value exceeds the threshold, the information is transmitted to the output layer.

![Figure 1 BP neural network structure diagram](image)

The classical structure of a BP neural network usually consists of an input layer, hidden layer, and output layer, each layer consists of multiple neurons, as shown in Figure 1. As the signal propagates forward, the neural network's predictions are output. As the signal propagates backward, the neural network learns by adjusting the weights and biases according to the direction of the fastest descent of the loss function gradient. A large number of experiments show that the three-layer BP neural network is very suitable for the regression analysis of the prediction model\(^5\), so this paper chooses the three-layer BP network for analysis.

The specific steps of the model establishment are as follows: (1) Data set preprocessing: normalize the collected original data set; (2) Construct PB neural network structure: determine the hidden layer of the network, generally, three layers, set the input layer, hidden layer, and output layer are the number of neurons; (3) BP neural network training; (4) BP neural network simulation prediction: test the prediction effect of the model on the training data and data set, and evaluate the effect of the output results, reliability is assessed.

The training set of the BP neural network generally includes forward and backward propagation parts. In the feedforward stage, the input layer neurons transmit the input pattern to the hidden layer. After calculating the weighted sum of each neuron in the output layer, the weighted sum is obtained through its activation function as the output value of the network. The error obtained in the feedforward stage of the neural network is transmitted backward from the output layer to the input layer. To reduce the network error to an acceptable level, the weights are rescaled according to the learning strategy.

1) Data Analysis and Preprocessing

When analyzing problems in real life, the original data we obtained is affected by the external environment, human subjective judgment, and other factors, which makes some data abnormal, missing, and other problems, so it may not be possible to obtain effective results in empirical research. in conclusion. Therefore, after obtaining the original data, the data must first be
preprocessed. The data in this paper is mainly aggregated by human calculation, and there may be data anomalies and some data missing, so it is necessary to preprocess the sample data. In this paper, data outlier correction, missing value completion, and data normalization preprocessing are performed on the data.

a) Boxplot method

Before processing the outliers of the original data, we must first judge whether they obey the normal distribution, and then choose the appropriate processing method. The original data was imported into SPSS24 for descriptive statistical analysis, and the results found that it did not obey the normal distribution. Therefore, according to the type of data, this paper selects the boxplot to preprocess the data.

The box plot was proposed by American statistician John Tukey, also known as box plot or boxplot, which can be used to observe the distribution and center position of sample data. The boxplot can be very intuitive to distinguish the outliers of the original data, and can also judge the discrete degree of the original data set by the length of the box and the length of the whiskers. The boxplot contains 5 numerical points from top to bottom, namely the upper whisker (Min), lower quartile (Q1), median (Median), upper quartile (Q3), and lower whisker (Max). Among them, the calculation formulas of the upper and lower whiskers are $Q3 + 1.5 \times (Q3 - Q1)$ and $Q1 - 1.5 \times (Q3 - Q1)$, the upper and lower quartiles and the median are the values $Q1$, $Q3$, and $Q2$ corresponding to the 25%, 75%, and 50% quartiles, respectively. The definition range of outliers in the original data set can be calculated by $[Q1 - 1.5 \times IQR, Q3 + 1.5 \times IQR]$. The 1.5 in the formula is the constant value $K$ assigned in this paper concerning the general practice, and whether each value is this or not is calculated by Python. If it is not within the range, it is an outlier in the test data set. The outliers in the original sample data can be screened out by using the boxplot method for the next step of data correction.

b) Data normalization processing

Data normalization refers to the need to normalize the data to make the data have the same measurement standard when the subsets or indicators in the data set have different characteristic value ranges. The data in the normalized dataset has the same limited range, which eliminates the influence of different dimensions between different data and improves the accuracy of the empirical results. According to the requirements of this paper, a neural network prediction model needs to be established, so the general practice is to normalize the data first, and this paper adopts the most valuable normalization method to process the data.

The most value normalization is to map all the data in the data set within the value range of $(0, 1)$, the two parallel lines represent the data set before and after mapping, The maximum mapping value is 1, the minimum mapping value is 0, and the calculation formula of the maximum value normalization method is as follows (2):

$$X_{scale} = \frac{X - x_{min}}{x_{max} - x_{min}}$$  

where $X_{max}$ and $X_{min}$ refer to the maximum and minimum boundaries of the metadata set data. $X$ refers to the value before normalization, and $X_{scale}$ represents the value after normalization.
c) Missing value handling

There are many ways to deal with missing values: mean interpolation, maximum likelihood estimation, multiple imputations, etc. However, for the processing of missing values, it is necessary to identify whether the attribute of the data set is cross-sectional data or time series data, and different data completion methods should be adopted for different types of data. It can be observed that the data obtained in this paper belong to time series data, so the direct deletion of missing data, the nearest supplementary method, the mean interpolation method, or the linear interpolation method can be used. In this paper, the linear interpolation method is used to carry out the missing data. The linear interpolation method assumes in advance that the time series has an obvious trend of change, and fills in the missing values by fitting the changing trend of the data.

2) BP neural network model test

In this paper, mean square errors (MSEs) and coefficient of determination (R2) are used to evaluate the predictive ability of network models. The MSE is calculated to represent the error of the BP neural network during training and testing, and its calculation formula is as follows. MSE represents the degree of deviation between the predicted value and the monitored value. The closer the MSE is to 0, the more accurate the predicted value. R2 was used to evaluate the goodness of fit of the network model to the dataset. The closer R2 is to 1, the higher the degree of explanation of the input variable to the output variable, and the better the prediction effect of the model. The formula for the coefficient of determination is expressed as follows (3) and (4).

\[ \text{MSE} = \frac{1}{2n} \sum_{i=1}^{n} f^2 (\sigma_i) \]  

\[ R^2 = \frac{\sum_{i=1}^{n} (y_i - \bar{y})^2}{\sum_{i=1}^{n} (y_i - \hat{y})^2} \]  

Among them, \( y_i \) is the true value; \( \hat{y}_i \) is the predicted value; \( \bar{y} \) is the average value of the true value; \( n \) is the number of samples in the training set or test set.

a) Mean Squared Error

The mean error refers to the expected value of the difference between the predicted output data of the data set and the target output data. The lower the value of MSE, the better the model effect is. At 0, there is no error. In addition, the x-axis in the figure represents the number of times the BP neural network is trained, and the y-axis represents the mean squared error value. The position of the coordinate point marked by the green circle in the figure represents the number of iterations of the neural network and the size of the value when the validation set is the best mean square error value. It can be seen from Figure 2 that there is an inflection point at (1, 0.02), indicating that the model is reasonable.
b) R coefficient
The R coefficient refers to the correlation between the predicted output data of the data set and the target output data. The closer the value of R is to 1, the closer the relationship between the prediction and the output data is, and the closer the value of R is to 0, it represents the prediction and the relationship between the output data is more random. The following four graphs represent the data dependencies of the training set, validation set, test set, and overall results after training, respectively. The abscissa represents the target output, and the ordinate represents the fitting function between the predicted output and the target output. It can be seen from Figure 3 that the correlation coefficients are 0.99074, 0.9999, 0.99596, and 0.98225, which are all close to 1, indicating that the fitting effect is good.

2) Analysis based on BP neural network results
According to the daily changes in the number of new local infections and new asymptomatic infections in the whole city and various districts of the new crown epidemic in Changchun since the outbreak, and after data cleaning, we obtained the data from March 29 to April 12 in
Changchun City. The number of vegetable bags distributed by the district. To know whether the supply of vegetable bags can reduce the contact of people and thus achieve the effect of suppressing the epidemic, this paper adopts the method of BP neural network, assuming that there is no supply of vegetable bags from March 29 to April 12, the number of people infected with the new crown make predictions. Before doing this research, we expected that if the distribution of vegetable packs would have an inhibitory effect on the epidemic, then the daily growth of the number of new crown infections would appear faster than the "V" peak, that is, the occurrence cycle of new crown infections would be shortened. Based on this conjecture, we have obtained the actual and predicted growth trend of the new crown infection as shown in Figure 4. It can be seen from the figure that the development trend of the epidemic reached a small peak on April 7 under the assumption that there is no supply of vegetable packs, and then began to decline. On the premise of the supply of vegetable packs, the growth rate of the number of new crown infections peaked around April 1 and has been on a downward trend since then. Therefore, it can be determined that the distribution of vegetable packs in Changchun City during the epidemic period can play a certain role in the control of the epidemic.

![Figure 4](image)

**Figure 4** The growth trend of the actual and predicted number of new crown infections

### 3 CONCLUSION

To explore the relationship between the control of the epidemic and the distribution of living materials, the development of the epidemic was mainly judged by the impact of the epidemic before and after the distribution of vegetable packs. Based on this research purpose, we conducted a PERSON correlation analysis on the number of COVID-19 infections and the number of vegetable bags distributed after vegetable bags were distributed in various districts of Changchun City, and concluded that vegetable bags would have an impact on the control of the epidemic[7]. To further verify, the article uses the BP neural network model to predict the development trend of the number of new crown infections in the future under the premise that there is no distribution of vegetable bags. Therefore, this paper concludes that the scientific management and distribution of living materials during the epidemic can introduce intensive contact between people, reduce risks, and achieve the effect of inhibiting the development of the epidemic.
REFERENCES


Research on Service Quality of College Express Market from the Perspective of Customers - Take Universities in Nanchang as an Example

Rui Zhang*
*Corresponding author: zrui199818@163.com
Beijing Jiaotong University, China

Abstract: The rapid development of e-commerce and the increase of national income have promoted the rapid development of the express industry. Many enterprises have entered the express market, making the express industry market competitive and complex. How to achieve long-term development in the fierce market competition depends largely on the quality of express service it provides. As the main force of online shopping, college students' loyalty to express enterprises is an important economic support for express enterprises. With the continuous expansion of college students' online shopping market, college express services have gradually revealed some problems, such as: loss of goods, inconvenience in returning goods, etc. Especially during e-commerce activities, the rapid growth of express business has brought greater challenges to express enterprises to ensure good logistics service quality. Therefore, it is very important to analyze the service quality of the efficient express market. This paper takes colleges and universities in Changbei Economic Development Zone of Nanchang City as the research object. From the perspective of customer perception, with reference to SERVQUAL model, LSQ model and other research results, this paper analyzes the factors that affect the service quality of college express market, and improves the college express service mechanism to improve customer satisfaction, expand and stabilize the college market.

Keywords: Logistics Service Quality, Logistics Service Quality, Empirical Research.

1 INTRODUCTION

In recent years, with the rapid popularization of online shopping, the business volume of China's express logistics industry has shown a rapid growth. From 2010 to 2019, the total business volume of express industry has kept growing year by year. In 2019, the business volume of express industry completed 63.52 billion pieces, with a year-on-year growth of 25.3%. Although the growth rate slowed down compared with previous years, it still maintained a rapid growth rate. As of 2020, the business volume of express service enterprises has completed 83.36 billion pieces, up 31.2% year on year. However, with the continuous diversification of customer needs, the existing express service quality can’t meet customer requirements. Especially in the period of e-commerce promotion activities, the volume of express delivery business increased sharply, and enterprises were unable to guarantee the quality of logistics services. A series of problems such as delay, damage and missing of express delivery appeared, which made the number of complaints or appeals about express
delivery services increase rapidly. Some scholars pointed out that good service quality can have a positive impact on customer satisfaction, and it is also closely related to consumers' purchase decisions and purchase intentions. In addition, more and more customers focus on the improvement of fast logistics service quality, and take the express service quality as the evaluation index of satisfaction. Therefore, how to develop new customers and maintain the existing customer stickiness in the fierce market competition, enhance the competitiveness of enterprises, and achieve long-term development is an urgent problem for express service enterprises to solve.

The most classic dimension division of logistics service quality is the SERVQUAL scale developed by PZB (1988). PZB initially proposed a gap model. On this basis, it further studied the service quality evaluation and proposed several factors that affect the logistics service quality, including a total of 97 test items. Later, it was reduced to 22 items and 5 dimensions were tangibility, reliability, responsiveness, assurance and empathy, Finally, a widely used SERVQUAL scale was formed. Zhang et al. (2002) divided the logistics service quality into five dimensions: service practicability, reliability, responsiveness, insurance and persistence by analyzing the traditional logistics service indicator system. Zheng et al. (2007) divided the quality of logistics service into time quality, personnel communication quality, order completion quality, coordination processing quality, flexibility, excellent product quality and convenience. From the perspective of customers, the LSQ model chart was constructed by adopting in-depth interview technology and Delphi expert opinion method, forming the evaluation indicators of China's local logistics service quality, and empirical testing was carried out. Wang et al. (2015)[9], based on SERVOUQUAL model and LSQ evaluation model, referred to the evaluation methods established by relevant scholars, further combined the characteristics of B2C e-commerce industry and the principle of B2C e-commerce logistics services, improved the original evaluation dimensions, and finally determined that convenience, responsiveness, reliability, empathy and information are the five dimensions to measure the quality of B2C e-commerce logistics services. Xu et al. (2009) proposed an e-commerce express service quality evaluation system consisting of four dimensions (empathy, responsiveness, tangibility and reliability) and 20 indicators (delivery time, price charges, etc.), and conducted a comprehensive analysis of the factors affecting the express service quality. In the study of Lan et al. (2017), express delivery service quality was defined: after customers experienced the services provided by express service organizations, the result of experiencing the service was express delivery service quality, and the content of express logistics service quality was defined from the four perspectives of security, timeliness, accuracy and convenience. Liu et al. (2010)[7] studied the relationship among logistics service quality, customer trust and use intention under the e-commerce environment based on the E-TRUST trust table. The research shows that the information quality, after-sales level and distribution quality of logistics services have a significant and positive impact on customer trust and use intention. Dai et al. (2014) studied customer satisfaction of logistics service in B2C environment. The results show that perceived reliability, perceived politeness and other service quality factors have a significant effect on customer satisfaction, and customer satisfaction has a significant effect on customer loyalty. Liao et al (2016)[10] took reliability, quality of service personnel, service flexibility, timeliness of distribution, information quality and economy as logistics service evaluation dimensions from the perspective of customer satisfaction. Its essence is to use SERVPREF model to evaluate logistics service quality. Zhan
et al. (2017), Zhang et al. (2016), Wang et al. (2019) discussed the influencing factors of campus logistics service quality from five aspects.

From the perspective of customer perception, based on the evaluation of express service quality by campus customers, this paper studies the service quality of campus express in Changbei University, providing reference for campus express companies to improve customer satisfaction, helping express companies effectively solve the problems of campus express, and meeting the needs of university teachers and students for express service. The rest of this paper mainly includes: theory and research hypothesis, methodology, analysis and results, as well as conclusions and suggestions.

2 THEORY AND RESEARCH HYPOTHESES

2.1 Identification of Variables

The collection and selection of literature follow the following five main steps:

(1) Classification background: The background of literature classification is the relationship between the effectiveness evaluation index system of logistics service quality and the logistics service quality and customer satisfaction.

(2) Analysis unit: the analysis unit is a scientific paper, including academic journal papers, working papers, etc.

(3) Literature collection: search for keywords from CNKI database, Baidu Academic, ScienceDirect and other English databases, and search for keywords and strings in publication titles and abstracts (such as "logistics service quality", "logistics service", "customer satisfaction", "customer perception", "empirical research" and their combinations).

(4) Scope definition: Without considering duplication, the result of the previous step is a group of 513 qualified papers, which are then screened according to specific criteria. Through extensive search, it is found that a large number of papers mainly focus on the logistics service quality optimization or logistics service quality evaluation model of an industry. Therefore, this paper first reviews the abstracts of relevant literature to ensure that the central theme is relevant. Papers whose topics do not meet the expected scope of this study are excluded. Finally, 55 papers published since 1988 were selected for in-depth review.

Referring to the existing research results, combined with the current situation of the express logistics service in Changbei universities, this paper summarizes the factors that affect the
quality of the campus express logistics service as five first level indicators: convenience, responsiveness, reliability, care and economy, and 15 second level indicators: convenience of sending and picking up, convenience of information query, and payment method; The specific description of each secondary indicator is shown in Table 1.

Table 1. Connotation of Factors Affecting the Quality of Express Logistics Service in Colleges and Universities

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Convenience</td>
<td>Convenience of sending and picking up</td>
<td>Whether it is convenient to pick up, whether it supports multiple ways of sending, and whether the sending process is complicated</td>
</tr>
<tr>
<td></td>
<td>Information query convenience</td>
<td>Whether it can be queried in real time and the query way is simple</td>
</tr>
<tr>
<td></td>
<td>Diversity of payment methods</td>
<td>Whether express fee supports multiple payment methods</td>
</tr>
<tr>
<td></td>
<td>Rationality of express site setting</td>
<td>Is the distance from the express point to the dormitory appropriate</td>
</tr>
<tr>
<td></td>
<td>environment of express site</td>
<td>Whether the site of express outlets is clean; Whether shelf placement and channel setting are convenient for taking parts</td>
</tr>
<tr>
<td>Responsiveness</td>
<td>Responsiveness of complaint handling</td>
<td>Can you quickly respond to customer complaints or feedback</td>
</tr>
<tr>
<td></td>
<td>Responsiveness of question answering</td>
<td>Can you quickly answer the customer's questions</td>
</tr>
<tr>
<td></td>
<td>Responsiveness of sending and picking up</td>
<td>Whether it can quickly respond to the customer's demand for delivery and pickup</td>
</tr>
<tr>
<td>Reliability</td>
<td>The express delivery has no missing or damage</td>
<td>The outer package and internal articles of express delivery shall be free from damage, loss or shortage</td>
</tr>
<tr>
<td></td>
<td>The logistics information is correct</td>
<td>The queried logistics information is consistent with the actual logistics information</td>
</tr>
<tr>
<td></td>
<td>Privacy information encryption</td>
<td>Whether the phone name or article information of the recipient or sender is encrypted</td>
</tr>
<tr>
<td>Caring</td>
<td>Service attitude of express station staff</td>
<td>The service staff had a good attitude and friendly communication with customers during the service process</td>
</tr>
<tr>
<td></td>
<td>Service level of express station personnel</td>
<td>Whether the service personnel are familiar with the operation process, whether the packaging is professional, and whether they can provide fast and complete services; Whether violent sorting exists</td>
</tr>
<tr>
<td>Economy</td>
<td>The price of the delivery</td>
<td>Whether the delivery price is reasonable</td>
</tr>
<tr>
<td></td>
<td>Additional service price</td>
<td>Whether additional payment is required when additional services are selected for delivery or pickup, and whether the payment is reasonable</td>
</tr>
</tbody>
</table>
2.2 Conceptual Model

The conceptual model of this study is shown in Figure 2. On the basis of previous literature, five factor groups that directly affect customer satisfaction with logistics service quality are determined and assumed, namely convenience, responsiveness, reliability, care and economy.

![Conceptual Model Diagram]

We assume the following:

**H1**: Logistics service convenience positively affects customer satisfaction with logistics service quality;  
**H2**: Responsiveness of logistics service positively affects customer satisfaction with logistics service quality;  
**H3**: The reliability of logistics service positively affects customer satisfaction with logistics service quality;  
**H4**: The caring nature of logistics services positively affects customers’ satisfaction with logistics service quality;  
**H5**: The economy of logistics service positively affects customer satisfaction with logistics service quality;

3 METHODOLOGY

3.1 Sampling and Data Collection

![Research Steps Diagram]
In order to design the questionnaire, the key variables and existing classifications of the research subject were determined through the literature review method, which has the following advantages: through the literature review method, we can comprehensively understand the current situation and problems of college express logistics service, the satisfaction of college student group customers with service quality, the factors affecting satisfaction and the research framework of related topics; The research limitations in the relevant literature can be studied by associating with other relevant literature. In this study, the five dimensions of convenience, responsiveness, reliability, care and economy of college express logistics service are identified as the key variables for the study, because after literature review, it is found that these five dimensions can well measure the quality of college express logistics service, and are also the key factors affecting customer satisfaction with logistics service. However, these potential variables cannot directly reflect the relationship between customer satisfaction and logistics service quality. Therefore, we continue to classify these potential variables based on literature reading and practical research so that they can be better used.

Based on the previous work, the questionnaire was designed and put into use.

3.2 Research Methods

The purpose of this study is to reflect the quality level of express logistics service in Changbei University and its impact on customer satisfaction through the five dimensions of convenience, responsiveness, reliability, care and economy of university express logistics service. Based on the following stages, determine the key variables and evaluate the specific process of their impact on express logistics service quality: conduct exploratory research according to the common characteristics of the identified variables to determine the influencing factors. Secondly, a confirmatory study is used to test the validity of the hypothesis model. Once the model is validated, the structural equation modeling method is used to test the proposed research hypothesis.

4 ANALYSIS AND RESULTS

4.1 Exploratory Study

Factor analysis is a common statistical method to analyze the relationship among a group of variables according to their common potential factors. The exploratory analysis in this paper is mainly carried out in three steps to check the adequacy of samples, extract key factors and test the reliability of the structure.

First, KMO is used to test the adequacy of data sampling and Bartlett's ball test is used to evaluate whether the variables determined by the team are suitable for factor analysis. KMO and Bartlett test results are shown in Table 2

<table>
<thead>
<tr>
<th>Table 2 KMO and Bartlett's test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kaiser Meyer Olkin measurement of sampling adequacy</td>
</tr>
<tr>
<td>Bartlett’s Sphericity Test</td>
</tr>
</tbody>
</table>
It can be seen from Table 2 that the KMO test result is 0.875, greater than 0.5. It can be seen that there is no significant difference in the correlation between variables, indicating that this sample collection is sufficient; The significance level of Bartlett's spherical test was 0.000, p<0.005, indicating that there was correlation between the original variables, which was suitable for factor analysis. KMO and Bartley's spherical test results both indicate that the data collected in this questionnaire are suitable for factor analysis.

Calculate the Kronbach coefficient of the overall sample to verify the reliability of the variable structure. The Kronbach coefficient of the overall sample is 0.881, which is greater than 0.7, indicating that the variable structure of the group study has good reliability.

### 4.2 Confirmatory Study

Amos was used for confirmatory factor analysis (CFA) to assess the adequacy of the goodness of fit between the basic factor structure and the sample data. Various fitting indexes can be used to test the overall model fitting, including goodness of fit index (GFI), adjusted goodness of fit index (AGFI), comparison fitting index (CFI), standard fitting index (NFI), non-standard fitting index (NNFI) and root mean square approximation error (RMSEA). Table 3 gives the appropriate model fit (χ² (190)=293.44, RMSEA=0.046, CFI=0.962, GFI=0.921, AGFI=0.902, NFI=0.907), which indicates that the model is completely acceptable. These estimates from the factor model show powerful results from the dataset. And all χ²The difference was significant (p<0.01), which supported the discriminant validity. The comprehensive reliability value of learning interaction input, learning emotional input, learning cognitive input and learning behavior input was 0.702, 0.701, 0.703, 0.730. All reliability estimates exceed the acceptable level of 0.7. The estimated AVE values of the five factor groups were 0.563, 0.610, 0.582 and 0.600 respectively, exceeding the critical value of 0.50.

#### Table 3 Confirmatory study results

<table>
<thead>
<tr>
<th>Latent variable</th>
<th>AVE</th>
<th>CR</th>
<th>Adaptability index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Convenience</td>
<td>0.5563366667</td>
<td>0.702715064</td>
<td>GFI=0.921; AGFI=0.902; IFI=0.963; RFI=0.932; NFI=0.907; CFI=0.962; RMSEA=0.046; X²/df=2.32;</td>
</tr>
<tr>
<td>Responsiveness</td>
<td>0.60895</td>
<td>0.700788308</td>
<td></td>
</tr>
<tr>
<td>Reliability</td>
<td>0.581716667</td>
<td>0.703646957</td>
<td></td>
</tr>
<tr>
<td>Caring</td>
<td>0.60014</td>
<td>0.7229383</td>
<td></td>
</tr>
<tr>
<td>Economy</td>
<td>0.58601</td>
<td>0.702250614</td>
<td></td>
</tr>
</tbody>
</table>

### 4.3 Structural Model and Hypothesis Testing

The structural model is used to test the assumptions of the four factors tested in the measurement model. In this study, SEM and maximum likelihood estimation methods are used to estimate the causal relationship between factors. Through the application of Amos to a
sample of 238 students, the analysis results are obtained, and the impact of these factors on the quality of college express logistics service is revealed. As shown in Table 4, the model fitting index is $\chi^2=413.06$, CFI=0.94, NNFI=0.92, RMSEA=0.052, standard $\chi^2=2.174$, which indicates that the model is fully consistent with the data. Therefore, the model is considered acceptable.

<table>
<thead>
<tr>
<th>hypothesis</th>
<th>Path coefficient</th>
<th>t value</th>
<th>testing results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Convenience→Logistics quality</td>
<td>0.7</td>
<td>3.38</td>
<td>Accept the original assumption</td>
</tr>
<tr>
<td>Responsiveness→Logistics service quality</td>
<td>0.62</td>
<td>2.96</td>
<td>Accept the original assumption</td>
</tr>
<tr>
<td>Reliability→Logistics service quality</td>
<td>0.61</td>
<td>2.87</td>
<td>Accept the original assumption</td>
</tr>
<tr>
<td>Caring→Logistics service quality</td>
<td>0.54</td>
<td>2.21</td>
<td>Accept the original assumption</td>
</tr>
<tr>
<td>Economy→Logistics service quality</td>
<td>0.57</td>
<td>2.43</td>
<td>Accept the original assumption</td>
</tr>
</tbody>
</table>

5 CONCLUSION AND DISCUSSION

In order to study how to better improve the quality of college express logistics service and improve customer satisfaction, this paper develops a model to evaluate the impact of convenience, responsiveness, care, reliability and economy on college express logistics service quality using SEM. The data of this paper mainly comes from 238 valid questionnaires collected from the Changbei University.

In general, convenience, responsiveness, caring, reliability and economy are all positively related to logistics service quality and customer satisfaction. The relationship between convenience, logistics service quality and customer satisfaction is the most significant, followed by responsiveness and reliability; Poor caring performance.

According to the research conclusions of this paper, this paper puts forward the following suggestions on how to better improve the quality of express service:

(1) Rectify the network environment and support multiple delivery and payment methods. The storage network of the school has a small space, and packages are crowded and disordered on the shelves. The network should reasonably place the shelves and set channels to facilitate picking up according to its actual business needs; At the same time, outlets should support multiple delivery methods and payment methods. Delivery methods include door-to-door delivery, door-to-door pickup and self-service delivery;

(2) Broaden service complaint channels. For problems that can not be solved in time and complaints can not be handled in time, express outlets can set up complaint mailboxes in the store and build customer WeChat groups online to keep communication with customers at any time to ensure timely receipt of customer feedback. The complaint information shall be cleared every day, and the complaint of the same day shall be handled on the same day, so as to maximize the customer service experience.
REFERENCES

Promoting Class Safety Benchmarking Evaluation Based on Pearson Correlation Analysis

Xiuhua Li¹,a*, Lei Fu²,b, Chao Zhang³,c, Shangmin Li¹,d, Yao Zhang¹,e, Jianing Si¹,f
fallingsnow201007@163.com¹, 296217107@qq.com², 17767665759@163.com³,
18813123217@163.com⁴, 15689709495@163.com⁵, JianingSi@hotmail.com⁶
State Grid of China Technology College, Jinan, China¹
Shandong Jining Tobacco Co. LTD, Jining, China²
State Grid Xinjiang Electric Power Co., Ltd, Bazhou Power Supply Company, Korla, China³

Abstract: With the proliferation of information technology, big data analysis has found wide applications in today’s safety management. Based on a Pearson correlation analysis between the incidence of unsafe student behavior and the coverage of unsafe student behavior by class safety benchmarking evaluation, this study focuses on how to reduce the incidence of unsafe student behaviors by promoting class safety benchmarking evaluation. The results show that promoting class safety benchmarking evaluation improves the rate of supervision on unsafe student behaviors and significantly reduces the incidence of unsafe student behaviors.

Keywords: Class Safety Benchmarking Evaluation, Correlation Analysis, Unsafe Student Behavior.

1 INTRODUCTION

The safety of college students has become an urgent issue as higher vocational colleges expand their enrollment. According to some research, college students, who are active in thinking but relatively weak in awareness of safety precautions, sometimes conduct unsafe behaviors that are in violation of school disciplines, healthy lifestyles, ethics or even laws, and some of the unsafe behaviors lead to campus safety incidents such as negative public opinion events, accidental injuries, etc., seriously endangering campus security and stability as well as students healthy growth [1, 3, 8]. Therefore, reducing the incidence of unsafe student behaviors is a most important task for colleges.

In today’s information age, big data has been widely used in safety management, with the basic function to figure out the laws for the incidence of unsafe behaviors from massive data, improve the safety supervision ability and effectively reduce unsafe behaviors [11]. In the study, the correlation between the incidence of unsafe student behaviors and the coverage of unsafe student behaviors by class benchmarking evaluation was evaluated, by retrieving big data on unsafe student behaviors from Campus B of D College for Pearson Correlation Coefficient Analysis. Based on the correlation analysis, the class safety benchmarking evaluation is promoted to improve the rate of supervision on unsafe student behaviors, with the goal of significantly reducing the incidence of unsafe student behaviors.
2 THEORETICAL BASIS FOR BENCHMARKING EVALUATION

As to how to reduce the incidence of unsafe student behaviors, scholars have conducted much research on the aspects of safety warning education, mental health services, safety management improvement, etc. \[1, 4, 6\]. While this study focuses on how to reduce the incidence of unsafe behaviors among students by improving the supervision of unsafe behaviors through correlation analysis-based class safety benchmarking evaluation.

Benchmarking, originating in the United States in the 1970s, can be defined as "the continuous process of analyzing the gap with best practice and taking measures to achieve fundamental improvement and innovation", and it has the essential feature in the pursuit of competitive advantage in modern knowledge management \[9\]. For optimizing benchmarking management, an important segment is implementing benchmarking evaluation to encourage the formation of shared responsibility \[7\]. In college, the class is the basic unit of safety management and class safety benchmarking is a process of promoting the class to keep a competitive state in safety management \[5\]. Similarly, for the smooth development of class safety benchmarking, a basic guarantee is promoting class safety benchmarking evaluation, by which the safety performance evaluation of each class member is integrated with the performance evaluation of the class as well as of the associated class cadres, with the aim of organizing the class members to form a community with shared responsibility. In this study, class safety benchmarking evaluation was promoted to reduce the incidence of unsafe student behaviors based on Pearson’s correlation analysis between the incidence of unsafe student behaviors and the coverage of unsafe student behaviors by class safety benchmarking evaluation.

3 CORRELATION ANALYSIS

For this study, we took 32 classes of 1636 freshmen from Campus B of D College as the subjects and retrieved big data on their unsafe behaviors. Through surveys, tests and correlation analysis of big data, we calculated the current incidence of unsafe student behaviors, the coverage of unsafe student behaviors by class benchmarking evaluation and the Pearson correlation coefficient between the two.

3.1 Current incidence of unsafe student behaviors

The incidence of unsafe student behaviors, an indicator of the density of unsafe behaviors, is the ratio of the number of unsafe student behaviors to the total number of students in a given period. To ensure campus safety and stability as well as students’ healthy growth, Campus B of D College proposed a requirement to reduce the incidence of unsafe behaviors among students to less than 6% by benchmarking against the historical best level of 5.83% and the current level of 5.97% of peer Campus A. By reviewing student unsafe behavior inspection records from October to December 2021, we calculated that the average monthly incidence of unsafe student behaviors was 9.71% with an interval of 9.53-9.83%, missing the target of 6%.
3.2 Current coverage of unsafe student behaviors by class benchmarking evaluation

According to the current rules on the evaluation of unsafe student behavior enforced by Campus B, all unsafe behaviors are counted in the performance evaluation of individuals, but only seriously disciplinary unsafe behaviors, such as getting drunk and staying out late, are counted in the performance evaluation of the class and its associated cadres. We looked at the evaluation records from October to December 2021 and calculated that the coverage of unsafe behaviors by class safety benchmarking evaluation was 13.2% on Campus B, which is much lower than the 61.2% on peer Campus A, as shown in Table 1.

<table>
<thead>
<tr>
<th>Campus</th>
<th>Number of unsafe behaviors</th>
<th>Number of unsafe behaviors covered by benchmarking evaluation</th>
<th>Coverage of unsafe behaviors by benchmarking evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Campus B</td>
<td>477</td>
<td>63 (in violation of disciplines that may lead to serious consequences)</td>
<td>13.2%</td>
</tr>
<tr>
<td>Campus A</td>
<td>281</td>
<td>172 (that may lead to serious consequences)</td>
<td>61.2%</td>
</tr>
</tbody>
</table>

We further examined the record of unsafe behavior supervision by student cadres under the current evaluation by taking as a sample the unsafe behavior of not measuring the temperature on time. As shown in Fig. 1 below, the average rate of supervision on non-punctual temperature measurements by associated class cadres was 51.2%, with an interval of 46.5-57.3%, failing to reach 80%, which is the eligible rate of supervision on unsafe conduct as specified in the cadre liability list.

![Figure 1: Rate of supervision on unsafe behaviors by associated class cadres](image)

In summary, currently the coverage of unsafe student behaviors by class benchmarking evaluation on Campus B is only 13.2%, much lower than 61.2% on peer Campus A. In this case, the rate of supervision on unsafe behavior by associated class cadres is lower than the value specified in the cadre liability list.
3.3 Correlation between the incidence of unsafe behaviors and the coverage of unsafe behaviors by class benchmarking evaluation

To further analyze the correlation between the incidence of unsafe behaviors and the coverage of unsafe behaviors by class benchmarking evaluation, we selected 32 classes and divided them into four experimental groups for the evaluation system transition test, controlling the other variables as much as possible. The groupings of the evaluation transition tests are listed in Table 2 below.

Table 2: Experimental groups for the evaluation transition test

<table>
<thead>
<tr>
<th>Groups</th>
<th>Classes</th>
<th>Coverage of unsafe behaviors by class benchmarking evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>Classes 1-8</td>
<td>Only unsafe behaviors in serious violation of disciplines are covered.</td>
</tr>
<tr>
<td>Group 2</td>
<td>Classes 9-16</td>
<td>2 common unsafe behaviors (not turning off the light on time and not measuring the temperature on time) are added.</td>
</tr>
<tr>
<td>Group 3</td>
<td>Classes 17-24</td>
<td>4 common unsafe behaviors (not turning off the light on time, not measuring the temperature on time, irregular use of electricity and not cleaning up the dormitory as required) are added.</td>
</tr>
<tr>
<td>Group 4</td>
<td>Classes 25-32</td>
<td>All unsafe behaviors except accidental ones are covered.</td>
</tr>
</tbody>
</table>

By referring to the records of safety inspection, unsafe behavior supervision and benchmarking evaluation of 32 classes in March, we calculated the coverage of unsafe behaviors by class safety benchmarking evaluation, the rate of supervision on unsafe behaviors by class cadres and the incidence of unsafe behaviors. As shown in Fig. 2, the coverage of unsafe behavior by class safety benchmarking evaluation, positively correlated with the rate of supervision on unsafe behaviors by class cadres, is negatively correlated with the incidence of unsafe behaviors.

![Figure 2: Results of evaluation transition test](image)

The Pearson Correlation Coefficient Calculator was used to calculate the correlation coefficient between the coverage of unsafe behavior by Class benchmarking evaluation and the incidence of unsafe behaviors for each class of the 4 groups in March. As can be seen from the XY scatter plot in Fig. 3, R has a value of -0.9652 and R² has a value of 0.9316, suggesting a strong negative
correlation between the coverage of unsafe behaviors by class benchmarking evaluation and the incidence of unsafe behaviors.

Figure 3: Correlation coefficient between the coverage of unsafe behaviors by class benchmarking evaluation and the incidence of unsafe behaviors

In addition, as shown in the XY scatter plot in Fig. 4 below, the incidence of unsafe behaviors is no longer decreased significantly when the coverage of unsafe behavior by class benchmarking evaluation rises above 90%.

Figure 4: The XY scatter plot of the coverage of unsafe behaviors by Class benchmarking evaluation and the incidence of unsafe behaviors
4 COUNTERMEASURES TO PROMOTE CLASS BENCHMARKING EVALUATION

Based on the above correlation analysis, classroom benchmarking evaluation is promoted to reduce the incidence of unsafe student behaviors to less than 6%, by increasing the coverage of unsafe behaviors by class benchmarking evaluation to 90% and the rate of supervision on unsafe behaviors by associated class cadres to 80%.

4.1 Formulating rules for class safety benchmarking evaluation

Taking the selection of excellent classes as an opportunity, we formulate rules for class benchmarking evaluation with a combination of rewards and penalties, as shown in Fig. 5 below. With multiple-evaluation of the relevant individuals, the associated student cadres and the class in two ways of "reward and penalty" for unsafe behaviors, the rules are formulated to build a new safety management ecology in which mutual supervision coexists with mutual achievement.

![Figure 5: Rules for class benchmarking evaluation](image)

4.2 Designing procedures for class benchmarking evaluation

Referring to the management process of students' monthly comprehensive quality evaluation, we designed the procedures of "daily inspection publicity, weekly summary notification, monthly statistical score deduction, and periodical evaluation and reward" for class benchmarking evaluation.

4.3 Promoting the implementation of class benchmarking evaluation

First, at the beginning of April, the rules for class benchmarking evaluation were released for all classes, thereby motivating the students to strive for excellence. Next, the class benchmarking evaluation was performed, and each unsafe behavior was also factored into the performance evaluation of the class cadres proportionally, as well as into the selection of good classes. Finally, the ranking of each class in the class benchmarking was computed inversely proportional to the evaluation of each class member. In addition, an additional quota of
outstanding cadres was given to the selected elite classes to inspire a sense of collective honor and responsibility, as well as vigor in supervision over unsafe behavior.

4.4 Driving the improvement in benchmarking evaluation

Class benchmarking evaluation is not a short-term behavior, but "a dynamic management process of continuous improvement" \[10\]. In order to continuously improve the class benchmarking evaluation, we timely assess the impact of class benchmarking evaluation. First, we conducted an online collection of student proposals for class benchmarking evaluation. For example, we actively responded to student voices by awarding each student who has never engaged in unsafe behavior an additional 1 point. Second, we accumulated and solidified the experience to assess the system by classifying the benchmarking evaluation cases. Third, based on a survey of student attitudes towards evaluation, we formulated improvement plans and continuously improve the class benchmarking evaluation, forming a closed loop of benchmarking, summarization and upgrading.

5 RESULTS OF CLASS BENCHMARKING EVALUATION

Class benchmarking evaluation was conducted from April to June. Then, by referring to the records of safety evaluation and conducting questionnaire surveys, we counted the coverage of unsafe behaviors by class benchmarking evaluation, the rate of supervision for unsafe behavior by associated class cadres and the incidence of unsafe behaviors, to check the implementation effect of class benchmarking evaluation.

5.1 Coverage of unsafe behaviors by class benchmarking evaluation

According to the safety evaluation record, 273 out of 282 unsafe behaviors conducted by individual students from April to June, with the exception of 9 accidental unsafe behaviors, were counted proportionally in the performance evaluation of student cadres, as well as in the selection of excellent classes. The coverage of unsafe behaviors by class safety benchmarking evaluation reached 96.8%, above the target of 90%. At the same time, the ranking of each class in the class benchmarking was computed conversely to the performance evaluation of each class member proportionally. In addition, for each of the top 10 selected outstanding classes, an additional outstanding cadre quota was awarded to the class cadre committee.

5.2 Rate of supervision on unsafe behavior by class cadres

We selected the unsafe behavior of "not measuring temperature on time" which is easy to be recorded as the sample and investigated the rate of supervision on unsafe behaviors by class cadres before and after the implementation of "multiple two-way" evaluation, as shown in Fig. 6 below. After the implementation of class benchmarking evaluation, the average rate of supervision on unsafe behavior by class cadres reached 94.8%, which is above the target value of 80%.
5.3 Incidence of unsafe behaviors

The implementation of class benchmarking evaluations was completed in July 2021 and continued from August to December 2021. Based on the normalized safety inspection data for October to December 2021, the incidence of unsafe behaviors among students decreased to 5.63%, below the target value of 6%. By comparing the number of unsafe behaviors for October to December 2021 with those for October to December 2020, it can be seen that the number of serious unsafe behaviors decreased from 48 to 11, and the number of safety emergencies such as accidental injuries decreased from 4 to 2, which decreased by 336% and 100% respectively.

6 CONCLUSIONS

It has been shown that by promoting class benchmarking evaluation, the incidence of unsafe behaviors fell below the target value of 6% and the rate of supervision on unsafe behaviors by class cadres rose above 80%. This study preliminarily explores how to reduce students’ unsafe behaviors through benchmarking evaluation, which has reference significance for higher vocational colleges creating a safety management situation in which the students are encouraged to grow into high-quality skilled talents by enhancing their awareness of safety, sense of responsibility and spirit of teamwork. But how to optimize the safety benchmarking evaluation index system to better leverage the role of benchmarking evaluation in improving students’ safety literacy remains a topic for further research.

REFERENCES

Literature Review of The Impact of China-ASEAN Free Trade Area on Industrial Structure Based on Citespace
-- Concurrently Discussing the Impact of RCEP on China's Industrial Structure Upgrading

Shuimeng Wan1, *, Xinzao Huang2
*ylyh06@163.com1, 005079@jxnu.edu.cn2

Department of International Education, Jiangxi Normal University, Master Degree Candidate, China1
The Regional Development Research Institute, Jiangxi Normal University, Research Supervisor, China2

Abstract: The China-ASEAN Free Trade Area (FTA), as the first FTA constructed in China, is an important initiative for China's industries to participate in the global division of labour and cooperation. This paper adopts the bibliometric analysis software citespace from the technical level and summarizes the existing impact of China-ASEAN free trade agreement on the upgrading of China's industrial structure and its path, and prospects the opportunities and challenges brought by RCEP to China's industrial structure. Specifically, based on the CiteSpace metrology analysis software based on Citespace analysis, the paper searches the literature on China-ASEAN FTA from two major databases, CNKI and Web of since, to sort out the research trends on China-ASEAN FTA and summarize the industrial clustering and industrial transfer effects of China-ASEAN FTA. Secondly, the path of the impact of the China-ASEAN FTA on the industrial structure is summarized, and the general view is that the trade effect as well as the tariff effect of the FTA will promote the optimisation of the industrial structure. Finally, we analyse how the upgraded FTA agreement RCEP will bring opportunities and challenges for China's industrial structure upgrading.

Keywords: China-ASEAN Free Trade Area, Industrial Structure, Citespace, RCEP.

1 INTRODUCTION

The 17th and 18th National Congresses of the Party clearly pointed out the need to accelerate the implementation of the free trade zone strategy, and the report of the 20th National Congress of the Party once again reiterated the importance of the free trade zone to China's economic development. The China-ASEAN Free Trade Area is the earliest, fastest progressing and most effective FTA in China[1]. According to the General Administration of Customs of China, the total import and export between China and ASEAN was only US$40 billion in 2003, while this figure jumped to US$878.2 billion in 2021, with an average annual growth rate of 37.3%. Even in the face of repeated shocks such as the financial crisis and public health incidents, export trade between the two sides has continued to grow steadily and even against the trend, showing strong vitality and dynamism[2]. Industry is the mainstay of economic
development, and the upgrading and transformation of industrial structure and opening up to the outside world are the core elements of building an FTA\(^3\). The result of the counter-trend growth of trade between the two sides cannot be separated from the support of industrial structure. Therefore, the relationship between the China-ASEAN FTA and China's industrial structure has been an important issue of concern to scholars. By selecting relevant representative research literature, this paper condenses the following three aspects: firstly, to summarise the general research trend of China-ASEAN FTA and analyse its impact on China's industrial structure; secondly, the path analysis of the impact of China-ASEAN FTA on China's industrial structure; thirdly, the opportunities and challenges of China's industrial structure under RCEP Opportunities and Challenges of Upgrading.

2 CHINA-ASEAN FREE TRADE AREA RESEARCH TRENDS AND INDUSTRIAL EFFECTS

2.1 Analysis of Research Hotspots in the China-ASEAN Free Trade Area

The China-ASEAN trade area is the first negotiated trade area in China, and its wide coverage and far-reaching influence have made it a focus and hot spot for academic research for a long time.

In this paper, through a search of the CNKI database, we found a total of 1677 papers related to the China-ASEAN Free Trade Area, including 210 core journals and 172 degree theses. After eliminating the literature unrelated to trade or industry, 225 pieces of literature remained. Research was still in its infancy at the early stage of the construction of the China-ASEAN FTA, and the research fever reached its peak in 2013-2017, and the research fever regarding the China-ASEAN FTA has remained enduring in recent years. From the keyword mapping at home and abroad (as shown in Figure 1), among the 225 sample documents in the CNKI database, industrial structure, industrial cooperation and industrial upgrading recurred more frequently and became the focus of scholars' attention. From Figure 2, it is easy to see that the trend of foreign research on FTAs is more empirical, generally based on the gravity model to analyse trade creation and trade transfer effects.

![Figure 1](image1) Key word map of the free trade zone
From the analysis of keyword emergence graphs at home and abroad (as shown in Figure 2), there are fewer studies on the industrial structure of the China-ASEAN FTA, and studies about industrial upgrading have only been paid attention to in recent years. Research trends in previous years abroad have concentrated on the gravitational model of the FTA, and studies on liberalisation and economic growth have only been given attention by scholars in recent years.

**Figure 2** Comparative analysis of domestic and international research trends on FTAs

### 2.2 The Industrial Effects of the FTA

#### 2.2.1 Industrial Transfer Effects

Japanese economist Akamatsu was the first to conduct a systematic study on the effects of industrial transfer. Based on him, Kiyoshi Kojima proposed the theory of marginal industrial expansion, emphasising that industrial transfer should transfer industries in which the country does not have a comparative advantage to other countries.

Taking the China-ASEAN FTA as an example, Qin (2005) divided the industrial development levels of member countries in the FTA into four gradations, pointing out that countries with higher industrial levels would transfer their inferior industries to countries with lower levels, so as to optimise their industrial layout[4]. After the establishment of the China-ASEAN FTA, tariff barriers have been removed and the exchange of goods and the flow of factors within the region has become freer, which has had an accelerating effect on the transfer of industries within the region[5].

#### 2.2.2 Industrial Agglomeration Effect

The process of China-ASEAN FTA leading to the transfer of industries is also accompanied by industrial agglomeration effects.

Wang Zhanbo et al. (2013) point out that the construction of the China-ASEAN FTA will result in the transfer of industries from countries within the region to countries outside the region, resulting in the phenomenon of industrial agglomeration and leading to uneven industrial development[6]. Long, Yunan (2013), also based on the exploration of industrial
agglomeration paths, reaches the opposite conclusion. He finds that the China-ASEAN FTA not only highlights the agglomeration effect, but also moderates the imbalance in industrial agglomeration[7]. Scholars generally agree that the China-ASEAN FTA brings about industrial agglomeration, but further evidence is needed to determine whether industrial agglomeration causes industrial imbalances.

3 MECHANISMS OF THE IMPACT OF THE CHINA-ASEAN FTA ON CHINA'S INDUSTRIAL STRUCTURE

3.1 Pathway Exploration

3.1.1 Trade Effects

Through literature analysis found that the keyword clustering as shown in the figure (as shown in Figure 3), get 323 connection node, found in the literature about China-asean free trade agreement industry structure, generally trade effect as the intermediary mechanism, think free trade area will stimulate the import and export goods and trade structure adjustment, ultimately affect the industrial structure.

Cui Qingbo et al. (2017) found that the establishment of the China-ASEAN FTA will promote trade liberalization between China and ASEAN countries and is conducive to China's international industrial upgrading and transformation[8]. Xiang Yan (2022) measured the trade creation and trade transfer effects of the China-ASEAN FTA based on a gravity model and found that both sides lacked comparative advantages in capital and technology-intensive industries and that such industries were difficult to be replaced[9].

3.1.2 Tariff Effects

Tariff reduction is a central part of the FTA negotiation and construction process, so scholars at home and abroad inevitably take the economic effects of tariffs as a starting point when studying the construction of FTAs, and delve into the chain reaction under the zero-tariff conditions in FTAs.

Cao, Liang et al. (2022) point out that tariff reductions allow enterprises to reduce the cost of importing intermediate goods from ASEAN, reduce the technical complexity of their exports,
and thus promote high-quality agricultural development. Kwanyoung et al. (2018) dissect the effect of tariff concessions in the China-ASEAN FTA based on the theory of customs union, and the study shows that tariff concessions can promote trade exchanges among member countries. Yu Miaojie and Wang Xiaotong (2021) show that tariff concessions in the China-ASEAN FTA lead to the entry of efficient firms instead of inefficient ones, which in turn leads to an increase in the overall productivity of the industry. Both the increase in production efficiency and the increase in trade volume are important factors affecting the industry, however, scholars have only skimmed over the research on industrial structure while studying the tariff effects of the FTA, failing to place the various effects on the upgrading of industrial structure and ignoring the impact of the various effects on industrial structure.

3.2 Empirical Study

In terms of ex-ante forecasting, global trade analysis model (GTAP model) is widely used to evaluate the effect of international trade policy on economy. This model evolved from general equilibrium model (CGE model) and has been widely used in the of trade liberalization and regional economic integration. Post-hoc assessment uses the trade gravity model to evaluate the effect of the China-ASEAN Free Trade Area. Roberts (2004) used the macroscopic gravity model to analyze the trade effect of China-ASEAN Free Trade Area, The establishment of the FTA led to a significant increase in trade flows between China and ASEAN. In addition to the above methods, some scholars have also used other methods to measure the chain reaction brought about by the FTA. Liang Shuanglu et al. (2020) compared the China-ASEAN FTA with the European and US FTAs and found that the two FTAs have different impact paths on industrial structure, with the China-ASEAN FTA mainly optimising the investment structure to achieve industrial upgrading.

4 THE IMPACT OF RCEP ON OUR INDUSTRIAL STRUCTURE

4.1 Opportunities for China's industrial Structure Arising from RCEP

At a time when countries are facing the double test of the new crown epidemic and the global economic downturn, RCEP will bring significant trade effects and accelerate the restructuring of the regional industrial chain layout. Wang Jianfeng et al. (2022) found that RCEP not only improves export trade efficiency, but also accelerates structural reforms on the supply and demand sides, adding momentum to industrial chain collaboration. The signing of RCEP will expand the trade creation effect, optimise resource allocation and promote deeper integration of industrial and value chains in the region.

In addition, RCEP can reconfigure the industrial chain and accelerate the reconfiguration of the regional industrial chain layout. Yali Wang et al. (2021) explain that RCEP provides a supply-side basis for intra-regional value chain climbing, with RCEP member countries gaining comparative advantage at low cost and becoming global production network centres. The implementation of RCEP brings opportunities to Chinese industries, and
seizing this opportunity can enhance the effective competitiveness of industries and strengthen their ability to withstand risks.

4.2 RCEP Challenges to Our Industrial Structure

While there is no doubt that the entry into force of the RCEP brings opportunities for restructuring industrial chains, some scholars are concerned about the risks and challenges that the RCEP brings to the industrial structure.

RCEP affects commodity import and export trade, which in turn changes the trade structure and puts invisible pressure on industrial adjustment. Yuan Minglan and Zhang Xiaoling (2021) argue that the similarity in export structures between the two sides has intensified competition between manufacturing and agriculture, putting pressure on industrial chain reconfiguration. Of course, there are also scholars who see the potential threats of the RCEP from other perspectives. The most direct manifestation of this game is the high wall of technical barriers that member countries have erected against each other, which puts technology in a jam and inhibits the development of high-tech industries. The tendency of capital to profit will transfer the low-end domestic manufacturing industry to Southeast Asian countries with lower labour costs, which means that China's industry will face the pressure of industrial hollowing out.

5 SUMMARY

From a theoretical perspective, the research system on the trade effect, tariff effect and industrial effect of the China-ASEAN FTA has tended to mature at home and abroad, but the current studies on the industrial structure of the China-ASEAN FTA are all based on the trade effect and tariff effect as the mediating mechanism, and we have yet to see a comprehensive study on the path of industrial upgrading in the China-ASEAN FTA. In terms of empirical research, studies on the trade effects of the FTA are mainly based on global trade analysis models and gravity models, which are widely adopted by academics, but lack direct measurement of industrial structure. Therefore, there is still room for research on the path of China-ASEAN FTA on the upgrading of industrial structure. After the introduction of the upgraded FTA RCEP, scholars have focused on the opportunities and challenges of the RCEP in building the industrial division of labour system and reconstructing the industrial chain, but have also neglected the exploration of the path of upgrading the industrial structure. After the launch of the upgraded China-ASEAN Free Trade Zone, namely RCEP, scholars have paid close attention to the opportunities and challenges of RCEP in building the industrial division of labor system and reconstructing the industrial chain, but also ignored the exploration of the path of industrial structure upgrading. Therefore, in the context of the redistribution of the global industrial chain, studying how to use RCEP to promote the upgrading of China's industrial structure will become a research hotspot of theoretical and practical significance.

REFERENCES

The Influence of the Change of Labor Relations on Willingness to Share in Sharing Economy Based on the Structural Equation Model

Junfeng Liao¹,a*, Yongyu Liao¹,b, Zhengyan Yang ¹,c and Siyuan Li²,d

¹ljf@scut.edu.cn, liaoyongyu20@163.com, yzyan2020@163.com, tonylsy@163.com

Department of Electronic Business, South China University of Technology, Guangzhou, Guangdong, China

Institute of Telecommunications Guangzhou Campus, Guangzhou, Guangdong, China

Abstract: This paper aims to explain how the changes in labor relations under sharing economy affect the behavior of sharers. Based on the structural equation model(SEM), SPSS and AMOS in this paper are used to conduct statistical analysis on collected data and verify the influence mechanism. This paper concludes that the perceived changes in labor relations cannot directly affect the willingness to share. The perceived changes in labor relations have a significant impact on relational psychological contract and transactional psychological contract respectively. The two-dimensional structure of psychological contract plays an intermediary role between perceived changes in labor relations and willingness to share and the transactional psychological contract has a more significant intermediary effect than relational psychological contract.

Keywords: Sharing Economy, Labor Relations, Psychological Contract, Willingness To Share, SEM.

1 INTRODUCTION

The emergence of sharing economy has triggered unprecedented changes in production relations and supply models, leading to the need for some theories to be revised and supplemented. In previous studies, the objects of research in the sharing economy are consumers, sharers, and sharing platforms, among which consumer-oriented research can be classified as consumer behavior and sharing platform-oriented research can be classified as organizational behavior. The role of sharers become special and the mode of sharing economy has prompted changes in the production relationship between the sharers’ role and the sharing platform.

For the sharing platform, the sharer is the consumer because the sharer is consuming the information resources of the sharing platform such as server bandwidth, customer service line and so on. For general consumers, the sharer represents the sharing platform and provides services for them, and the service attitude of the sharer will affect the overall evaluation of the sharing platform by consumers. From this perspective, the sharer can be regarded as a member of the sharing platform organization. However, the sharer owns the means of production, that is, the sharer enters the organization with the means of production, which is quite different
from the traditional production relationship. Therefore, the sharer is not an employee in the strict sense for the sharing platform. There will be a new cooperative labor mode between the sharer and the sharing platform. Under this labor mode, the sharer's willingness to share will affect the management and stability of the sharing platform.

The concept of "psychological contract" was put forward by Argyris in 1960. The psychological contract theory focuses on the very clear and fixed strong production relationship between organizations and employees. However, the production relationship in the sharing economy shows that there is no strict employment relationship between sharers and sharing platforms, but there is still a weak production relationship between them. Under the new mode of sharing economy, whether psychological contract will affect sharing willingness and whether the change of labor relations in sharing economy will affect sharing willingness through psychological contract has not been verified. Therefore, this paper takes the online car sharers as the research object, and studies the sharers' willingness to share and the influencing mechanism based on the structural equation model.

2 MATERIALS AND METHODS

2.1 Perceived Changes in Labor Relations and Willingness to Share

Ren and Wang (2019) discussed the new changes in labor relations under the background of sharing economy compared with the traditional economy. Hoffman et al. (2016) advocated rebuilding the trust between enterprises and employees through the "alliance" strategy, and building a labor relations coordination mechanism conducive to mutual trust, mutual investment and common benefit. From the research at home and abroad, it is not difficult to find that the labor relationship between online car-sharing enterprises and sharers is different from the traditional form. The variable "perceived changes in labor relations" can be used to reflect the brand-new relationship between the sharer and the car-sharing enterprise, and to express the sharer's perception. In addition, employee dissatisfaction often leads to different levels of resistance or resistance psychology. Compared with traditional labor relations, sharers have more ways to express their dissatisfaction, such as commenting and sharing. The most direct trade-off for this labor relationship by sharers is the decrease in willingness to share. Based on this, this paper puts forward the hypothesis:

H1: Perceived changes in labor relations have a significant positive impact on willingness to share.

2.2 Perceived Changes in Labor Relations and Psychological Contract

The psychological contract of teachers in applied education is different from those in other fields. All these phenomena corroborate from the side that the change of perceived labor relations has a significant impact on the psychological contract. With the deepening of the relationship between the sharer and the platform, the psychological contract is also constantly strengthened. In addition, the three-dimensional structure of psychological contract is invalid under the background of sharing economy, and its two-dimensional structure includes relationship dimension and transaction dimension. The former is a broad, long-term and open responsibility, which pays more attention to future development and emotional exchange and
is more subjective. The latter is a concrete and short-term responsibility, which pays more attention to economic components, that is, the exchange relationship between material and society, which is in line with the core of psychological contract[6]. Based on this, this paper puts forward hypothesis 2(H2) and hypothesis 3(H3).

H2-H3: Perceived changes in labor relations have a significant positive impact on relational psychological contract and transactional psychological contract.

2.3 Psychological Contract and Willingness to Share

Davenport and Prusak(1997) argued that willingness to share was related to people's personality, and some people were naturally willing to share and did not care about rewards. The organizational environment and the relationship between the organization and employees will have an impact on employees' psychology, which will lead to the fluctuation of employees' willingness to share. Lu and Chen(2012) found that relational psychological contract had the greatest influence on tacit knowledge sharing willingness and transactional psychological contract was negatively related to explicit/implicit knowledge sharing. For the sharing economy, economic benefits had been proven to be one of the driving factors for participation in sharing, and the sharing platform provided users with direct or indirect subsidies to improve their willingness to participate in sharing[4]. Based on previous studies, this paper puts forward hypothesis 4(H4) and hypothesis 5(H5).

H4-H5: Relational psychological contract and transactional psychological contract have significant positive effects on willingness to sharing.

It has been clarified in the previous section that perceived labor relationship change may have some effect on both willingness to share and relational and transactional psychological contract, and it is highly likely that relational and transactional psychological contract also have significant effects on willingness to share. That is, on the premise that all the above five hypotheses may be established, this study reasonably speculates that psychological contract can be used as an intermediary variable, so this paper puts forward hypothesis 6(H6) and hypothesis 7(H7).

H6- H7: There is a mediating effect of relational psychological contract and transactional psychological contract.

The theoretical model of this study is shown in Figure 1.

![Image](image.png)

**Figure 1: Theoretical model diagram.**
3 RESEARCH DESIGN

3.1 Variable Measurement

In this study, Didi drivers were selected as the investigation object, and Likert 5 subscale was adopted. The questionnaire is divided into four parts. The first part is to measure the driver's perception of the change of labor relations with Didi Company. The second part is the psychological contract part, referring to the scale of Rousseau(2001) and Li(2006). The third part is the willingness to share, modified from the scale used by Hamari(2015). The fourth part is the basic information of the respondents: gender, age, education level, and frequency of driving online taxi. The details of the scale were shown in table 1.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Scale</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived changes in labor relations</td>
<td>Q11: I think I am a regular employee of Didi Company.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q12: I think I am a member of Didi Company.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q13: I think Didi’s staff is my superior.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q14: I think other Didi drivers are my colleagues.</td>
<td></td>
</tr>
<tr>
<td>Relational psychological contract</td>
<td>Q21: Didi Company provides a good communication channel.</td>
<td>Rousseau(2001)</td>
</tr>
<tr>
<td></td>
<td>Q22: Didi Company often listens to my opinions.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q23: Didi Company doesn’t set a threshold to prevent me from going to other online car platforms.</td>
<td>Li(2006)</td>
</tr>
<tr>
<td></td>
<td>Q24: Didi’s policy of protecting drivers’ personal safety is effective.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q25: In front of the passengers, I will maintain the image of Didi Company.</td>
<td></td>
</tr>
<tr>
<td>Transactional psychological contract</td>
<td>Q31: The task reward provided by Didi Company is reasonable.</td>
<td>Rousseau(2001)</td>
</tr>
<tr>
<td></td>
<td>Q32: The task reward provided by Didi Company is stable.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q33: The income provided by Didi Company is satisfactory to me.</td>
<td>Li(2006)</td>
</tr>
<tr>
<td></td>
<td>Q34: The bonus offered by Didi Company is attractive.</td>
<td></td>
</tr>
<tr>
<td>Willingness to share</td>
<td>Q41: I am willing to share my spare car.</td>
<td>Hamari(2015)</td>
</tr>
<tr>
<td></td>
<td>Q42: If necessary, I will help my friends to be Didi drivers.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q43: I will continue to be a Didi driver.</td>
<td></td>
</tr>
</tbody>
</table>

3.2 Pretest and Data Sources

To ensure the validity of the questionnaire, a pretest was first conducted for this study, and questionnaires were distributed in the gathering communities of Didi drivers. The number of pretest questionnaires was 109, and the effective recovery rate was 90%. Among them, there were 107 male respondents, accounting for 98.2%, and 2 female respondents, accounting for 1.8%. This gender structure was consistent with the actual situation.
The results of pretest were shown in table 2, Cronbach's Alpha values of four variables: perceived labor relationship change, relational psychological contract, transactional psychological contract and willingness to share were 0.888, 0.825, 0.866 and 0.876 respectively. Each variable basically met the requirements without modification for the time being. However, after exploratory factor analysis, it was found that the answers to the third question, the fourth question and the fifth question of the relational psychological contract scale and the third question of the transactional psychological contract scale were scattered and could not express the research content, so they were deleted.

Table 2: The pre-test results

<table>
<thead>
<tr>
<th>Model variable</th>
<th>Title number</th>
<th>α</th>
<th>Whether it reaches the standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived changes in labor relations</td>
<td>Q11, Q12, Q13, Q14</td>
<td>0.888</td>
<td>yes</td>
</tr>
<tr>
<td>Relational psychological contract</td>
<td>Q21, Q22, Q23, Q24, Q25</td>
<td>0.825</td>
<td>yes</td>
</tr>
<tr>
<td>Transactional psychological contract</td>
<td>Q31, Q32, Q33, Q34</td>
<td>0.866</td>
<td>yes</td>
</tr>
<tr>
<td>Willingness to share</td>
<td>Q41, Q42, Q43</td>
<td>0.876</td>
<td>yes</td>
</tr>
</tbody>
</table>

The revised questionnaire was distributed again. The formal questionnaire took the form of network and offline, and finally 354 questionnaires were collected, of which 329 were valid, with an effective rate of 92.94%. In this study, SPSS and AMOS 21.0 was used as an analysis tool, and SEM structural equation was used to process the core data.

4 EMPIRICAL RESULTS

4.1 Reliability Analysis

Reliability refers to the reliability and consistency of measurement results. Huang(2009) and other scholars believed that it could measure the difference between the design of the content of the questionnaire and the actual views of the respondents. If a survey carries out the same test on the same subject several times, the result can explain whether the data has the same trend. There is a positive correlation between reliability and consistency.

Since this questionnaire adopted Likert 5 subscale, Cronbach's alpha coefficient was used to measure reliability. According to scholars Shao and Yang, the alpha coefficient should be better than 0.6. Above 0.9 is "excellent", around 0.8 is "very good", 0.7 is "moderate", and above 0.5 is acceptable. The generally accepted standard in recent years is α greater than 0.7. SPSS 22.0 was used to analyze the reliability of the sample, and the results were shown in Table 3.
Table 3: Reliability analysis results

<table>
<thead>
<tr>
<th>Variables</th>
<th>Title number</th>
<th>α</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived changes in labor relations</td>
<td>011</td>
<td>0.904</td>
</tr>
<tr>
<td></td>
<td>012</td>
<td>0.904</td>
</tr>
<tr>
<td></td>
<td>013</td>
<td>0.907</td>
</tr>
<tr>
<td></td>
<td>014</td>
<td>0.908</td>
</tr>
<tr>
<td>Relational psychological contract</td>
<td>021</td>
<td>0.908</td>
</tr>
<tr>
<td></td>
<td>022</td>
<td>0.909</td>
</tr>
<tr>
<td></td>
<td>023</td>
<td>0.925</td>
</tr>
<tr>
<td>Transactional psychological contract</td>
<td>031</td>
<td>0.903</td>
</tr>
<tr>
<td></td>
<td>032</td>
<td>0.906</td>
</tr>
<tr>
<td></td>
<td>033</td>
<td>0.907</td>
</tr>
<tr>
<td>Willingness to share</td>
<td>041</td>
<td>0.920</td>
</tr>
<tr>
<td></td>
<td>042</td>
<td>0.905</td>
</tr>
<tr>
<td></td>
<td>043</td>
<td>0.920</td>
</tr>
</tbody>
</table>

From the above table, we can see that Cronbach's α coefficients were all greater than 0.9, so it could be seen that the measurement results of the four dimensions in the scale of this study were acceptable.

4.2 Validity Analysis

This scale used SPSS 22.0 to analyze the structure validity of the collected data according to the methods mentioned above. The specific results were shown in table 4.

Table 4: Validity analysis results

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>KMO</th>
<th>Significance of Bartlett ball test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived changes in labor relations</td>
<td>1.000</td>
<td>0.801</td>
<td>.000 Sig.&lt; 0.001, the spherical hypothesis is rejected, and factor analysis can be done.</td>
</tr>
<tr>
<td></td>
<td>1.215</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.938</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.772</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relational Psychological contract</td>
<td>1.000</td>
<td>0.699</td>
<td>.000 Sig.&lt; 0.001, the spherical hypothesis is rejected, and factor analysis can be done.</td>
</tr>
<tr>
<td></td>
<td>1.254</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.040</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transitional Psychological contract</td>
<td>1.000</td>
<td>0.666</td>
<td>.000 Sig.&lt; 0.001, the spherical hypothesis is rejected, and factor analysis can be done.</td>
</tr>
<tr>
<td></td>
<td>1.115</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.150</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Willingness to share</td>
<td>1.090</td>
<td>0.712</td>
<td>.000 Sig.&lt; 0.001, the spherical hypothesis is rejected, and factor analysis can be done.</td>
</tr>
<tr>
<td></td>
<td>0.874</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.082</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Whole sample</td>
<td>0.924</td>
<td>0.000</td>
<td></td>
</tr>
</tbody>
</table>

According to the value of this table, it could be judged that it was suitable for factor analysis. The KMO values of the whole sample and each variable were all greater than or equal to 0.5.
Besides, the significant probability of Bartlett's spherical test statistical value was 0.000, less than 0.001, so the spherical hypothesis was rejected. It was proved that the correlation coefficient matrix of the factor was a non-unit matrix, and could extract the least factor while explaining most of the variance. Therefore, factor analysis could be done. The factor rotation matrix in the factor analysis show that the items under each variable were aggregated, and the items among different variables were separated, indicating that it had good structural validity.

4.3 Model Fitting Index

The results of model fitting index were shown in table 5. The absolute fitting index was mainly used to test the fitting degree between the model proposed in this paper and the sample data actually collected. Common measurement indicators including the ratio of chi-square to degree of freedom (X2/df), goodness of fit index (GFI), root mean square of approximate error (RMSEA), and adjusted goodness of fit index (AGFI) in this study were 3.059, 0.079, 0.919 and 0.88 respectively, which all met the standard level. The relative fitting index was to compare the differences between the benchmark model and the model in the research. Non-normal fit index (NNFI, namely TLI) and comparative fit index (CFI) were 0.921 and 0.938 respectively, so the model proposed in this study and the sample data collected were in line with the index from the overall fit.

The reduced fitting index was the degree of simplification of the model mentioned above. It could be seen from the table 5 that the indicators included: PNFI, PGFI and PCFI, all of which were above 0.5 and met the standard level. This show that the simplification degree of the model proposed in this paper was acceptable. To sum up, the indexes in the model test of this study were all within the reasonable standard range and it was suitable for subsequent analysis.

<table>
<thead>
<tr>
<th>Fitting index</th>
<th>Standard level</th>
<th>The present study model</th>
<th>Is the study consistent with</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolute fitting index</td>
<td>X2/df</td>
<td>Between 2.0 and 5.0</td>
<td>3.059</td>
</tr>
<tr>
<td></td>
<td>RMSEA</td>
<td>Between 0.05 and 0.08</td>
<td>0.079</td>
</tr>
<tr>
<td></td>
<td>GFI</td>
<td>&gt;0.80</td>
<td>0.919</td>
</tr>
<tr>
<td></td>
<td>AGFI</td>
<td>&gt;0.80</td>
<td>0.880</td>
</tr>
<tr>
<td>Relative fitting index</td>
<td>TLI</td>
<td>&gt;0.80</td>
<td>0.921</td>
</tr>
<tr>
<td></td>
<td>CFI</td>
<td>&gt;0.80</td>
<td>0.938</td>
</tr>
<tr>
<td>Reduced fitting index</td>
<td>PNFI</td>
<td>&gt;0.50</td>
<td>0.713</td>
</tr>
<tr>
<td></td>
<td>PGFI</td>
<td>&gt;0.50</td>
<td>0.588</td>
</tr>
<tr>
<td></td>
<td>PCFI</td>
<td>&gt;0.50</td>
<td>0.616</td>
</tr>
</tbody>
</table>

4.4 Hypothesis testing

AMOS 21.0 was used to conduct path analysis in this paper. Of the seven hypotheses in this research model, six hypotheses (H2, H3, H4, H5, H6 and H7) reached the significance level
and were supported, while one hypothesis (H1) was rejected because it did not reach the significance level. The path analysis results were shown in table 6 and figure 2.

Table 6: Sample hypothesis test results.

<table>
<thead>
<tr>
<th>Assumption</th>
<th>Path</th>
<th>Path coefficient</th>
<th>Significance</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>Perceived changes in labor relations have a significant impact on willingness to share.</td>
<td>-0.103</td>
<td>Not significant</td>
<td>Not Established</td>
</tr>
<tr>
<td>H2</td>
<td>Perceived changes in labor relations have a significant impact on relational psychological contract.</td>
<td>0.536</td>
<td>*** significant</td>
<td>Established</td>
</tr>
<tr>
<td>H3</td>
<td>Perceived changes in labor relations have a significant impact on transactional psychological contract.</td>
<td>0.716</td>
<td>*** significant</td>
<td>Established</td>
</tr>
<tr>
<td>H4</td>
<td>Relational psychological contract has a significant impact on willingness to share.</td>
<td>0.619</td>
<td>*** significant</td>
<td>Established</td>
</tr>
<tr>
<td>H5</td>
<td>Transactional psychological contract has a significant impact on willingness to share.</td>
<td>0.550</td>
<td>*** significant</td>
<td>Established</td>
</tr>
<tr>
<td>H6</td>
<td>There is a mediating effect of relational psychological contract.</td>
<td>Compare to H1, H2 and H4.</td>
<td></td>
<td>Established</td>
</tr>
<tr>
<td>H7</td>
<td>There is a mediating effect of transactional psychological contract.</td>
<td>Compare to H1, H3 and H5.</td>
<td></td>
<td>Established</td>
</tr>
</tbody>
</table>

Note: *** p<0.001, ** p<0.01, * p<0.05.

Figure 2: Model inspection diagram.

Specifically, in the model, perceived changes in labor relations had a significantly positive impact both on relational psychological contract and transactional psychological contract and the path coefficients were 0.536 and 0.716 respectively, so H2 and H3 were supported in this study. There was a positive correlation between two kinds of psychological contract and willingness to share and the coefficients were 0.619 and 0.550. The stronger the psychological contract, the higher willingness to share. Therefore, H4 and H5 were valid. Perceived changes in labor relations had no significant influence on the willingness to share so H1 was rejected. The perceived changes in labor relations could not directly affect the sharing intention, but through the two-dimensional structure of psychological contract with intermediary effect, in which transactional psychological contract has a more significant impact than relational psychological contract. H1 was rejected and H7 was supported.
5 CONCLUSIONS AND IMPLICATIONS

5.1 Research Conclusion

Hypothesis 1 is not tenable, which is reasonable. Didi emphasized the cooperative relationship with the sharers at the beginning of its business development. Many drivers could accept this brand-new labor relationship and shared their private cars, which show that the sharers' willingness to share was still high when they perceived the change of labor relations. Moreover, whether labor relations significantly affect willingness to share has not been confirmed in previous studies, suggesting that the relationship between the two may be weak or even negligible.

Hypothesis 2 and 3 are significant, which shows that the change of perceived labor relations has a significant impact on relational psychological contracts and transactional psychological contracts respectively. In the investigation and interview, it was also found that some sharers indicated the shift was attractive to them because of the greater degree of freedom in time and work under the sharing model. Therefore, hypotheses 2 and 3 are consistent with the theory and reality.

Hypothesis 4 and 5 are significant, but the path coefficient is weak. In the study of Chen(2012), the fit indicators of psychological contract and sharing willingness were good; however Zhao(2005) found that the psychological contract of civil servants would be different in the reform environment than before. Combining the results of several scholars and the research in this paper, it can be seen that both relational and transactional psychological contracts can have an impact on willingness to share.

Hypothesis 6 and 7 are both significant. The effect of perceived labor relationship change on willingness to share becomes significant after the introduction of relational and transactional psychological contracts as mediating variables, indicating that both relational and transactional psychological contracts can be used as fully mediating variables. Mapping to real life, that is, psychological suggestion and material reward. Psychological suggestion is embodied in that the online car-sharing enterprise uses some measures to make the sharer think that the online car-sharing enterprise regards him as a member of the company, creating an emotional connection between the sharer and the online car-sharing enterprise. The results in this study also indicate that material rewards are more attractive to the sharers than psychological hints.

5.2 Practical Inspiration

5.2.1 Strengthen the awareness of labor relations of sharers

It is suggested that online car rental companies need to clarify labor relations as much as possible, such as through electronic contracts. Once the labor relationship is confirmed, the sharers will have more trust in the online car-related enterprises, and they will be more inclined to formal employees and tend to be stable in psychological contracts, thus effectively improving the sharers' willingness to work and helping the online car-related enterprises to grow further.
5.2.2 Broaden information channels and improve material rewards

Abundant feedback channels are conducive to the stability and healthy positive construction of relational psychological contract. Positive psychological contract can reduce the negative emotions of sharers, and alleviate the tense relationship between sharers and online car companies. Once there is a conflict of interest, it can also effectively reduce the breach of contract or retaliatory extreme behavior of sharers.

This study found that most sharers pay more attention to material rewards, and the impact of transactional psychological contract on the willingness to share is much higher than that of relational psychological contract. In order to reduce negative behaviors that are not conducive to online car-sharing enterprises, the best countermeasure is to improve material rewards, which can enhance the stability of sharers.

5.2.3 Strengthen the circulation speed of information and enhance sense of participation of sharers

User operation is related to user experience, and network car companies need to further increase their investment in user operation. Especially when the most influential material rewards fluctuate, network car companies should comfort the sharers in official channels, large social platforms, media and other information circulation places in time, and try their best to reduce the impact on sharers' willingness to share in this way. On the other hand, they should strengthen the speed of information circulation, ensure the timeliness of information circulation in official channels, large social media platforms and other information circulation places, thus enhancing sharers' participation and their willingness to share.

Acknowledgements: Thanks for the support from Philosophy and Social Science Planning Project of Guangzhou (2018GZYB13).

REFERENCES

Analysis of Exchange Rate Fluctuations between RMB and US Dollar

Zilin Liu¹⁺*, Wentian Wang²⁺, Chunlei Zhao¹⁺

⁺737827170@qq.com, S21301072@ahu.edu.cn, 1770150920@qq.com

School of International Business and Economics, Shandong University of Finance and Economics, Shandong, Jinan, China¹
School of Social and Political Science, Anhui University, Anhui, Hefei, China²

Abstract: Under the background of rapid development of big data and artificial intelligence technology, mining the information implied in time series data has become an important step for all industries to make decisions. At the same time, using big data for data analysis is a new requirement for economic work in the current rapid development of information technology, which is also an important way to achieve exchange rate analysis. At the 2022 Central Economic Work Conference, it was pointed out that the RMB exchange rate should be basically stable at a reasonable and balanced level, and the financial stability guarantee system should be strengthened. This paper mainly takes the central parity rate of the RMB against the U.S. dollar expressed by the direct pricing method from January 5, 2015 to January 31, 2021 as the research object. Through the GARCH model, it studies and analyzes the changes in the fluctuation of the RMB against the U.S. dollar exchange rate. The results show that there is a “herd effect” in people's expectations of exchange rate changes. At the same time, it finds that GARCH (2,1) and T-GARCH (1,1) predict the results significantly, and puts forward relevant suggestions and countermeasures.

Keywords: Exchange Rate, Dollar, Big Data.

1 INTRODUCTION

Facing the complicated domestic and international environment, the report of the 20th National Congress of the Communist Party of China clearly proposed that we should adhere to a high level of opening up, promote high-quality development, and accelerate the construction of a new development pattern. To better serve the dual cycle strategy and give play the role of the exchange rate as an "automatic stabilizer" in macroeconomic stability and the balance of international payments, which is the key point of unblocked internal and external circulation, the Central Bank clearly declared that it would enhance the flexibility of the RMB exchange rate and promote the reform of the RMB exchange rate formation mechanism.

At the same time, with the deepening of China's exchange rate reform, exchange rate fluctuations not only reflect the changes in the value of the RMB, but also have a signal function that comprehensively reflects the changes in China's economy and foreign economic cycles. Therefore, from the perspective of time series, it plays an important role in macroeconomic regulation and foreign exchange management to capture the regularity of
exchange rate changes and predict their future changes. The US dollar is one of the most important currencies in the world. Studying the RMB exchange rate fluctuation against the US dollar has become an urgent task to control and effectively prevent exchange rate risks. Based on the analysis of the exchange rate change between RMB and USD, this paper establishes a regression model to predict the future trend of the exchange rate change.

2 LITERATURE REVIEW

The change in the exchange rate is similar to the movement of objects in physics. From the concept of statistics, academic circles use variance/standard deviation, conditional heteroscedasticity and other indicators to measure exchange rate elasticity as the first choice. Levy Yayati and Sturzenegger use standard deviation to measure the volatility of exchange rates of various countries. Based on the data of 183 countries in the world from 1974 to 2001, they calculated the mean value of the absolute value of the monthly rate of change of exchange rate, the standard deviation of the monthly rate of change, and the mean value of the absolute value of the net reserve change relative to the base currency ratio. They determined the exchange rate regimes of different countries according to the combination of three Vars and classified the floating. The three types of intermediate and fixed exchange rate regimes are assigned values of 1, 2 and 3 respectively to indicate the elastic level of the exchange rate [8]. Similarly, Mahradika also uses standard deviation to measure elasticity [9]. Cavoli and Rajan further dynamically measured the standard deviation (variance) of elasticity and used conditional volatility to measure exchange rate elasticity [2]. Bleaney and Tian used root mean square error to measure the level of exchange rate elasticity, and believed that the pegged exchange rate system had a lower root mean square error; The floating exchange rate system had a high root mean square error; The exchange rate fluctuation performance of managed floating exchange rate system is similar to that of independent floating exchange rate system [1]. Some scholars further set exchange rate change rate indicators. Shambaugh (2004) divided the exchange rate regimes of different countries into two categories: pegged and non-pegged, with 2% as the boundary [11]. Reinhart and Rogoff studied the parallel and dual market exchange rate data of 153 countries from 1946 to 2001 in a five-year window period. They classified the elasticity level into five categories and thirteen subcategories by using exchange rate changes and assigned values from 1 to 13 [10].

The "exchange rate elasticity" constructed from the concept of economics reflects the response mechanism of exchange rate changes to some economic activities. For example, the exchange rate elasticity of foreign exchange reserves constructed by Cavoli, Rajan, Combes, etc. Actually reflects the situation that the change of foreign exchange reserves to adjust the exchange rate will cause the change of exchange rate. Zhang Xiang and others constructed the exchange rate elasticity of money supply based on the Dornbusch model. The functioning of these mechanisms must meet a default premise: a country's foreign exchange market is highly developed and the part of the relevant indicators (such as foreign exchange reserves, monetary volume and interest rate changes) that causes exchange rate changes can be identified (Zhang 2014). The foreign exchange markets of most developing countries, including China, are in transition. Not only may there be market and nonmarket factors in exchange rate changes, but also the extent to which changes in corresponding currencies, interest rates and foreign exchange reserves cause exchange rate changes is opaque. Ignoring these facts, there will be a
dilemma similar to that described in Levy Yeyati and Sturzenegger's study that China and India cannot be classified because their exchange rates and foreign exchange reserves are almost stable during the sample period. Domestic and foreign scholars have conducted in-depth research on the driving factors of exchange rate fluctuations. Lee found that when a country switches to a floating exchange rate system, its exchange rate fluctuations will more reflect macro fundamental factors (Lee 2007). Dabrowski et al. (2014) found that there is a long-term cointegration relationship between the exchange rate and economic fundamentals, and short-term changes in fundamentals will also have a significant impact on the exchange rate \cite{5}. Chou used the variance decomposition method to measure the relative contribution of macro fundamentals in explaining exchange rate fluctuations, and found that the fundamentals have significant explanatory power for exchange rate changes \cite{5}. Xie and Chen proved that observable fundamentals are an important reason for exchange rate fluctuations \cite{12}.

It can be found that the literature at home and abroad is mainly based on the classic exchange rate determinants model, using empirical methods to test the correctness of the exchange rate theoretical model. From the perspective of research content, most studies mainly focus on the exchange rate differences between China and the United States, and test the influencing factors of exchange rate based on the VAR model or multiple linear regression. Based on predecessors, this paper will use the GARCH model to analyze the exchange rate fluctuations and adverse movements between China and the United States.

3 DATA SOURCE AND VAR SELECTION

In this paper, 754 observations are selected as the sample space of the central parity rate between RMB and USD expressed by the direct bid method from January 5, 2015, to January 31, 2021, and a series \( p \) is established based on this. Then, a logarithmic series \( r \) of the central parity rate between RMB and USD is constructed, a conditional heteroscedasticity model is established for the series \( r \), and its volatility is studied.

Data from the official website of the National Bureau of Statistics (http://www.safe.gov.cn) EViews8.0 is used for data processing.

4 ANALYSIS PROCESS

4.1 Establishing Logarithmic Series

Since time series data tends to show an exponethmic yield series is established for the data of the above selected Vars.

In this case, take the logarithm of the data of sequence \( p \) to obtain a new sequence \( r \).

The expression entered is: \( r = \log p \)
4.2 Draw Logarithmic Sequence Sequence Diagram

![Figure 1: Logarithmic sequence sequence diagram](image)

From the time series diagram of the logarithmic series \( r \) of the direct pricing method of RMB against USD, it can be observed that the fluctuation of the logarithmic series of RMB against the USD exchange rate is non-stationary. As shown in Figure 1.

4.3 Column Statistical Graph of Logarithmic Series

![Figure 2: Column statistical graph of logarithmic series](image)

From Figure 2, the mean value of the logarithmic series is 1.87681, the standard deviation is 0.039786, and the skewness is -0.344457, which is less than 0, indicating that the distribution of the logarithmic series has a long tail. The kurtosis is 2.02794, which is less than the kurtosis value 3 of the normal distribution, indicating that the logarithmic sequence has the characteristics of a thick tail.

4.4 Unit Root Test

Since the data used in this paper are only the results of logarithmic processing of objective data from statistical data, pseudo regression of non-stationary Vars may occur in the process of econometric analysis of these data. Therefore, a stationarity test should be conducted for these Vars.
With the help of EViews8.0 software, the unit root test was conducted on the logarithmic series \( r \) of the RMB US dollar exchange rate. The results are shown in Table 1:

**Table 1: Unit root test.**

<table>
<thead>
<tr>
<th>t-Statistic</th>
<th>Prob.*</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADF test statistic</td>
<td>-1.44</td>
</tr>
<tr>
<td>Test critical values:</td>
<td></td>
</tr>
<tr>
<td>1%level</td>
<td>-3.44</td>
</tr>
<tr>
<td>5%level</td>
<td>-2.87</td>
</tr>
<tr>
<td>10%level</td>
<td>-2.57</td>
</tr>
</tbody>
</table>

According to the ADF test results in Table 1, the t statistic is -1.438,773, and the corresponding P value is 0.5642. Therefore, it indicates that the logarithmic series \( r \) of the RMB US dollar exchange rate is non-stationary.

**Table 2: Unit Root Test after First Order Difference.**

<table>
<thead>
<tr>
<th>t-Statistic</th>
<th>Prob.*</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADF test statistic</td>
<td>-24.09</td>
</tr>
<tr>
<td>Test critical values:</td>
<td></td>
</tr>
<tr>
<td>1%level</td>
<td>-3.44</td>
</tr>
<tr>
<td>5%level</td>
<td>-2.87</td>
</tr>
<tr>
<td>10%level</td>
<td>-2.57</td>
</tr>
</tbody>
</table>


According to the ADF test results in Table 2, the t statistic is -24.09146, and the corresponding P value is close to 0. Therefore, it indicates that the first order difference of the logarithmic series \( r \) of the RMB dollar exchange rate is stable.

### 4.5 Sequence Autocorrelation and Partial Autocorrelation Test

**Figure 3** Sequence autocorrelation and partial autocorrelation test
It can be seen from Figure 3 that the autocorrelation Coef of the sequence shows the property of tailing, and the partial autocorrelation Coef shows the property of truncation after the first order, all falling within twice the estimated standard deviation. Therefore, the sequence determines that the model is AR (1).

<table>
<thead>
<tr>
<th>Var</th>
<th>Coef</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(1)</td>
<td>0.13</td>
<td>0.04</td>
<td>3.53</td>
<td>0.00</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.02</td>
<td>Mean dependent var</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.02</td>
<td>S.D.dependent var</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>S.E.</td>
<td>0.00</td>
<td>Akaike info criterion</td>
<td>-9.50</td>
<td></td>
</tr>
<tr>
<td>Res SS</td>
<td>0.00</td>
<td>Schwarz criterion</td>
<td>-9.49</td>
<td></td>
</tr>
<tr>
<td>Log likelihood</td>
<td>3571.26</td>
<td>Hannan-Quinn criter.</td>
<td>-9.49</td>
<td></td>
</tr>
<tr>
<td>D-Wn stat</td>
<td>2.01</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Build model:

$$D_r = 0.1278D_{r-1} + \epsilon_t$$

(3.5324)

$R^2 = 0.0159$ $\text{DW} = 2$

Pass the parameter significance test (as shown in Table 3), and check whether the residual error is white noise.

As shown in Figure 4, the results show that it is a white noise sequence with autocorrelation.
4.6 Perform the ARCH effect test.

<table>
<thead>
<tr>
<th>Table 4: ARCH Effect Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-statistic</td>
</tr>
<tr>
<td>Obs*R-squared</td>
</tr>
</tbody>
</table>

According to the results in Table 4 and Table 5, it can be found that the logarithmic series \( r \) of the RMB US dollar exchange rate has the ARCH effect, rejecting the original hypothesis (\( H_0: a_1=a_2=...=a_q=0 \), that is, unconditional heteroscedasticity effect).

<table>
<thead>
<tr>
<th>Table 5: RESID^2(-1) Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Var</td>
</tr>
<tr>
<td>C</td>
</tr>
<tr>
<td>RESID^2(-1)</td>
</tr>
<tr>
<td>R-squared</td>
</tr>
<tr>
<td>Adjusted R2</td>
</tr>
<tr>
<td>S.E.</td>
</tr>
<tr>
<td>Res SS</td>
</tr>
<tr>
<td>Log likelihood</td>
</tr>
<tr>
<td>F-statistic</td>
</tr>
<tr>
<td>Prob(F-statistic)</td>
</tr>
</tbody>
</table>

4.7 Perform the GARCH Effect Test

GARCH (1,1), GARCH (2,1) and GARCH (1,2) models were established respectively.

<table>
<thead>
<tr>
<th>Table 6: GARCH(1,1) Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Var</td>
</tr>
<tr>
<td>AR(1)</td>
</tr>
<tr>
<td>Variance Equation</td>
</tr>
<tr>
<td>C</td>
</tr>
<tr>
<td>RESID(-1)^2</td>
</tr>
<tr>
<td>GARCH(-1)</td>
</tr>
<tr>
<td>R-squared</td>
</tr>
<tr>
<td>Adjusted R2</td>
</tr>
<tr>
<td>S.E.</td>
</tr>
<tr>
<td>Res SS</td>
</tr>
<tr>
<td>Log likelihood</td>
</tr>
<tr>
<td>D-W stat</td>
</tr>
</tbody>
</table>
Table 7: GARCH(2,1) Model.

<table>
<thead>
<tr>
<th>Var</th>
<th>Coef</th>
<th>Std. Error</th>
<th>z-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(1)</td>
<td>0.17</td>
<td>0.03</td>
<td>5.06</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Variance Equation

<table>
<thead>
<tr>
<th>Var</th>
<th>Coef</th>
<th>Std. Error</th>
<th>z-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>0</td>
<td>0</td>
<td>5.99</td>
<td>0.00</td>
</tr>
<tr>
<td>RESID(-1)^2</td>
<td>0.34</td>
<td>0.06</td>
<td>5.85</td>
<td>0.00</td>
</tr>
<tr>
<td>RESID(-2)^2</td>
<td>-0.35</td>
<td>0.06</td>
<td>-5.94</td>
<td>0.00</td>
</tr>
<tr>
<td>GARCH(-1)</td>
<td>1.07</td>
<td>0.00</td>
<td>500.32</td>
<td>0.00</td>
</tr>
</tbody>
</table>

R-squared: 0.01  Mean dependent var: 0.00
Adjusted R2: 0.01  S.D. dependent var: 0.00
S.E.: 0.00  Akaike info criterion: -9.74
Res SS: 0.00  Schwarz criterion: -9.71
Log likelihood: 3668.54  Hannan-Quinn criter: -9.73
D-W stat: 2.09

Table 8: GARCH(1,2) Model.

<table>
<thead>
<tr>
<th>Var</th>
<th>Coef</th>
<th>Std. Error</th>
<th>z-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(1)</td>
<td>0.07</td>
<td>0.04</td>
<td>1.83</td>
<td>0.07</td>
</tr>
</tbody>
</table>

Variance Equation

<table>
<thead>
<tr>
<th>Var</th>
<th>Coef</th>
<th>Std. Error</th>
<th>z-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>0.00</td>
<td>0.00</td>
<td>5.57</td>
<td>0.00</td>
</tr>
<tr>
<td>RESID(-1)^2</td>
<td>0.15</td>
<td>0.03</td>
<td>5.06</td>
<td>0.00</td>
</tr>
<tr>
<td>GARCH(-1)</td>
<td>-0.09</td>
<td>0.02</td>
<td>-4.36</td>
<td>0.00</td>
</tr>
<tr>
<td>GARCH(-2)</td>
<td>0.74</td>
<td>0.05</td>
<td>14.83</td>
<td>0.00</td>
</tr>
</tbody>
</table>

R-squared: 0.01  Mean dependent var: 0.00
Adjusted R2: 0.01  S.D. dependent var: 0.00
S.E.: 0.00  Akaike info criterion: -9.63
Res SS: 0.00  Schwarz criterion: -9.61
Log likelihood: 3625.78  Hannan-Quinn criter: -9.62
D-W stat: 1.90

Based on the comparison of the above three models, it can be observed that all parameters of GARCH (2,1) pass the t-test with the best effect. As shown in Table 1, Table 2, and Table 3.

Then consider T-GARCH modeling. As shown Table 9:

Table 9: T-GARCH Model.

<table>
<thead>
<tr>
<th>Var</th>
<th>Coef</th>
<th>Std. Error</th>
<th>z-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(1)</td>
<td>0.09</td>
<td>0.04</td>
<td>1.94</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Variance Equation

<table>
<thead>
<tr>
<th>Var</th>
<th>Coef</th>
<th>Std. Error</th>
<th>z-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>0.00</td>
<td>0.00</td>
<td>7.11</td>
<td>0.00</td>
</tr>
</tbody>
</table>
Finally, the following equation is obtained:

\[ D_r = 0.088 \hat{D}_{r-1} + \varepsilon_t \]

\[ h_t = 1.26 + 0.952 \varepsilon_{t-1}^2 + 0.2779 \varepsilon_{t-2}^2 + 0.4946 \hat{h}_{t-1} \]

5 CONCLUSION

This paper takes the logarithmic series of the middle rate exchange rate under the direct pricing method of RMB against USD from January 5, 2015 to January 31, 2021 as the research object, and establishes a GARCH model for it, and draws the following conclusions:

First: Since 2015, the exchange rate of RMB against the US dollar has fluctuated greatly, showing a phenomenon of instability and stability after the first order difference, which is characterized by a thick tail. There is no autocorrelation, but there is significant conditional heteroscedasticity. It shows that there is a "herd effect" in people's expectations of exchange rate changes, that is, when the market expects the exchange rate to appreciate, more people expect the exchange rate to appreciate, and when the market expects the exchange rate to depreciate, more people expect the exchange rate to depreciate.

Second: From the perspective of the prediction ability of the model, GARCH (2,1) and T-GARCH (1,1) have the best prediction effect on the fluctuation of the US dollar exchange rate.

In this regard, we propose the following suggestions:

First: Study and legislate reasonable foreign exchange management policies and regulations. While continuing to promote the reform of the exchange rate formation mechanism, pay close attention to the changes in the ownership of RMB pricing right after the reform of foreign exchange and relevant influencing factors, maintain the gradual and controllable nature of the reform of foreign exchange, and strengthen the financial stability guarantee system.

Second: Establish a dynamic monitoring mechanism for changes in the frequency structure of RMB exchange rate fluctuations. In terms of policy orientation, we should continue to enhance the driving role of macro fundamentals in exchange rate fluctuations and spillovers,
maintain the basic stability of the RMB exchange rate at a reasonable and balanced level, and ensure that the impact of market factors on the RMB exchange rate is within a controllable range.

REFERENCES

Systematic Risk Prediction in Commercial Banks Based on Random Forest and BP Neural Network

Junbin Zhang¹,a, Peiying Zhang¹,b, Shiyang Song²,c, Junyu Su¹,d, Jinhai Tang²,e

f20092100174@cityu.mo, zpy1701@163.com,b, *Corresponding author: 2463756962@qq.com,c,
sjy2166@163.com,d, 1992649158@qq.come

Faculty of Finance, city University of Macau, Macau, China¹
Alibaba Cloud Big Data Application College, Zhuhai College of Science and Technology, Zhuhai, China²
Faculty of Data Science, City University of Macau, Macau, China³

Abstract: Since the 1990s, the frequent occurrence of systemic financial risks culminating in financial crises has had a serious impact on the economies and financial systems of all countries. Systemic risk analysis has become a very important task for most central banks in the wake of the global financial crisis (GFC). The sudden and destructive nature of systemic financial risks requires that we should pay attention to the foresight of systemic financial risks. In this study, based on establishing a system of systemic financial risk characteristics indicators in China, we construct machine learning models of random forest and support vector machine to warn systemic financial risks in China, and compare the warning effects of the two models using confusion matrix, ROC curve (Receiver Operating Characteristic Curve) and dynamic warning analysis, and The main factors that drive up the level of systemic financial risk in China are identified.

Keywords: Random Forest, BP Neural Network, Risk Forecast, Machine Learning.

1 INTRODUCTION

Because indirect financing dominates in China and the banking sector plays an integral role in the overall economic system, the banking sector is the key area of significant financial risk in China. It is inevitable that instability in the banking sector will result in significant losses for other economic institutions. The country's systemic financial risk prevention should be focused on preventing major risks to banks. A major role in early warning of systemic financial risks has traditionally been played by traditional statistical and measurement methods.

Nowadays, with the development of computer parallel computing power and data science, the frontier technology of big data has gradually penetrated into various fields and brought revolutionary changes to the modern financial industry. Big data analysis methods represented by machine learning and deep learning have a series of advantages such as timeliness, accuracy, and sample size, which make them well suited for data analysis and information processing in the financial field and further enrich the means and tools for systemic financial risk early warning [4].
In recessions, neuro-fuzzy models can improve the forecasting efficiency of daily stock market data, even though machine learning algorithms are still more effective during crisis periods [5]. A 45-year sample of banking systems in developed economies was analyzed to compare the out-of-sample forecasting ability of various early warning models, and machine learning algorithms were found to outperform Logit methods. A variety of complex patterns can be recognized by machine learning algorithms. In developing countries, machine learning algorithms can be used to enhance the efficiency of financial markets. The ability of machine learning algorithms to identify complex structures and make accurate predictions is attracting increasing attention from scholars. The need for intelligent risk warning systems is also high among industries and governments. For risk early warning research in financial markets, the use of machine learning algorithms will be crucial to building a risk control system.

Compared with the existing literature on systemic financial risk early warning, the contribution of this paper includes two main aspects: first, it adopts the frontier concept of "model uncertainty", and through a more standardized research model and analysis process, early warning of systemic financial risk of commercial banks is conducted based on big data methods such as random forest and BP neural network, and a better prediction model is found using model evaluation methods. [1]. Which improves the accuracy and effectiveness of systemic financial risk early warning in the banking industry to a large extent.

2 CURRENT RESEARCH STATUS

Systemic financial risk early warning techniques based on linear models are widely used in academia to detect economic crises and systemic financial risks. Research conducted by Chinese scholars using linear models such as FR has led to the establishment of an early warning system for major financial risks in China. To construct a capital market tail risk early warning system, an analysis of FR, STV, KLR and other early warning models was conducted [10]. In recent years, nonlinear models have replaced linear models in systematic financial risk forecasting and early warning. This is because linear models are unable to capture nonlinear relationships between economic and monetary variables, and they have enhanced early warning performance.

There has been an increasing amount of current research on the use of machine learning and deep learning nonlinear models to warn of systemic financial risks in foreign countries, and its analytical frameworks and analysis methods have also developed over time. In China, however, there is still relatively little research on the use of machine learning and deep learning to detect systemic financial risks early on. The purpose of this paper is to address the shortcomings of the existing literature by focusing on commercial banks and adopting machine learning and deep learning models for early warning analysis of systemic financial risks in China. The aim is to make early warning prediction of systemic financial risks more accurate and effective.

2.1 Random Forest

Random Forest is one of the more widely used and powerful machine learning methods, which is good at dealing with various types of prediction problems. Random forest is actually an integration of decision trees, which are usually trained by bagging method. Among them, a
decision tree is an algorithm that uses a tree structure to make decisions, consisting of a root node, several leaf nodes and internal nodes, where a leaf node represents a decision outcome and each other node represents an attribute test. A random forest is composed of several decision trees, and the best decision outcome is determined by voting on the decision outcome of each decision tree \(^2\). The basic idea of the random forest algorithm: k samples are drawn in the original training set using the bootstrap method with the same sample size as the original training set; k decision trees are built based on the k samples, resulting in k classification results; and each record in the k classification results is voted to determine its final classification.

2.2 BP Neural Network

BP neural network is a large nonlinear network, which simulates the human physiological reflex process. BP neural network simulates a large number of "neuron" nodes, in the input of a large amount of data, by itself to find the laws and logic between the nodes, and save the learning path between the nodes \(^1\). The concept of BP neural network was first introduced in the late 1980s, and its model building process is to explore the intrinsic connections and laws through the process of data input and output independently, without setting up the function relationship in advance. Through the gradient descent method, the BP neural network continuously adjusts its own weights, and when it does not reach a certain error accuracy, it reverses the stimulation until the final result meets the error accuracy.

The main idea of BP neural network is to estimate the error of the previous layer of the output layer based on the error after the output, and then use the error of this layer to estimate the error. The error of this layer is then used to estimate the error, so that the error estimates of all layers are obtained. The error estimate here can be understood as some kind of partial derivative, and we adjust the connection weights of each layer according to this partial derivative, and then recalculate the output error with the adjusted connection weights. BP neural network is a multilayer backward-looking intelligent network trained with error back propagation algorithm.

![BP neural network model diagram](image)

**Figure 1** BP neural network model diagram

2.3 Commercial Banking Systemic Risk

The Financial Stability Board defines systemic risk as the risk of experiencing a strong
systemic event, which can occur only as a result of a "systemic event". It is caused by the loss of a financial institution and the exposure of an institution with "systemic importance" characteristics, the impact of which on other institutions cannot be underestimated. Systemic risk arises from the activities of banks, and when the relationship between banks gradually forms a network, one bank is affected and the whole banking system is widely affected [8]. The problem of systemic financial risk is not limited to the economic and financial spheres of a country, even if a country with a modest economy, once financial turmoil is generated it often affects the economic and financial situation in other parts of the world. In the accumulation phase, systemic financial risks have one of the most important characteristics - they accumulate in the upward financial cycle, are insidious and not easily identified. The existing study [6] used machine learning techniques such as K-nearest neighbors, random forests, support vector machines, Boosting, and regulatory assessment data on bank risks to build a UK bank crisis early warning system and compared it with traditional statistical techniques such as logistic regression, which found that machine learning techniques significantly outperformed traditional statistical techniques such as logistic regression, and in particular, random forests performed well and were suitable as a bank algorithmic model of the crisis early warning system.

3 RESEARCH METHODOLOGY

3.1 Feature Selection and Index System Construction

In terms of feature selection, in order to control the influence of individual bank characteristics and macroeconomic variables on the systemic risk of banks, based on the existing literature, the paper selects 11 control variables to establish a systemic financial risk indicator system for China’s banking industry, taking into account the domestic and international research results and the actual national conditions of China.

<table>
<thead>
<tr>
<th>Variable Code</th>
<th>Variable Name</th>
<th>Variable Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>NI</td>
<td>Percentage of non-performing loans</td>
<td>Non-performing loans as a percentage of total loans</td>
</tr>
<tr>
<td>NM</td>
<td>Net income margin</td>
<td>Net income / Total assets</td>
</tr>
<tr>
<td>NP</td>
<td>Non-performing loans ratio</td>
<td>Non-performing loans / Total loans</td>
</tr>
<tr>
<td>ROE</td>
<td>Return on equity</td>
<td>Net income / Total equity</td>
</tr>
<tr>
<td>ROA</td>
<td>Return on assets</td>
<td>Net income / Total assets</td>
</tr>
<tr>
<td>MUE</td>
<td>Market value of equity</td>
<td>Market value of equity / Market value of equity</td>
</tr>
<tr>
<td>CDR</td>
<td>Capital adequacy ratio</td>
<td>Capital / Risk-weighted assets</td>
</tr>
<tr>
<td>SIZE</td>
<td>Size of bank</td>
<td>Total assets / Total assets</td>
</tr>
<tr>
<td>LEVERAGE</td>
<td>Leverage ratio</td>
<td>Total liabilities / Total assets</td>
</tr>
<tr>
<td>FD</td>
<td>Financial development</td>
<td>Stock market value / Total assets</td>
</tr>
<tr>
<td>MC_GR</td>
<td>Money growth rate</td>
<td>MC growth rate / GDP growth rate</td>
</tr>
<tr>
<td>GDP_GR</td>
<td>GDP growth rate</td>
<td>GDP growth rate / GDP growth rate</td>
</tr>
</tbody>
</table>
3.2 Model Performance Evaluation Methods

3.2.1 Confusion matrix and ROC curve

Confusion matrix and ROC curve to evaluate the goodness of a model requires performance metrics, i.e., designing evaluation criteria to measure the generalization ability of the model. In machine learning and deep learning classification tasks, confusion matrix and ROC curve are more commonly used performance measures \(^9\). Among them, the confusion matrix is a more comprehensive representation of the model evaluation results, which classifies the predicted samples based on whether the true values are the same as the predicted values. One row of the confusion matrix is used to represent the true category of the sample, and one column is used to represent the predicted category of the sample. When the predicted category of the sample is the same as the true category of the sample, it means that the model predicts the correct classification of the sample. When the sample prediction category is different from the sample true category, it means the model predicts the wrong classification of the sample. The ROC curve is a common tool for analyzing the classification behavior of models with different thresholds.

Table 2 Confusion matrix example

<table>
<thead>
<tr>
<th>True situation</th>
<th>Forecast result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk</td>
<td>TP (Real risk)</td>
</tr>
<tr>
<td>FP (False risk)</td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>FN (false normal)</td>
</tr>
<tr>
<td></td>
<td>TN (true normal)</td>
</tr>
</tbody>
</table>

3.2.2 Cross-validation

In recent years, cross-validation has been used to select model weights in various model settings, including heteroskedasticity linear regression models, linear regression models with lagged dependent variables. Cross-validation is a statistical method commonly used in machine learning and deep learning to evaluate the generalization performance of models through experimental tests, which is more scientific and reasonable than the usual method of dividing data into training and testing sets in a single pass \(^6\). In the cross-validation process, the data is generally partitioned several times and multiple models are trained at the same time. A common cross-validation method is k-fold cross-validation, where k is an arbitrary number that can be specified.

First, the data set is randomly cut into k disjoint subsets of the same size; then k-1 subsets are used as training sets to train the model, and the remaining (held out) one subset is used as a test set to test the model; the previous step is repeated for the possible k choices (each time a different subset is picked as the test set); thus k models are trained, and the test error is calculated for each model on the corresponding test set to obtain k test errors, and a cross-validation error is obtained by averaging these k test errors.
4 EXPERIMENTAL RESULTS

In this paper, we use data from 34 commercial banks in China, take 11 indicator variables in the systemic financial risk characteristic indicator system of China from January 2019 to December 2021 as the input of the early warning mode [3]. Take the sequence of risk early warning dummy variables obtained by transforming the results of risk monitoring analysis as the model expectation output, divide the data in a random single pass (75% of the data as the training set and 25% of the data as the test set), and the systematic financial risk early warning models of random forest and BP neural network are constructed in JupyterNotebook, which supports Python language, respectively. In this paper, 11 indicator variables in the systemic financial risk characteristic indicator system of China from January 2008-December 2017 are used as the input of the early warning model, and the sequence of risk warning dummy variables obtained by transforming the results of risk monitoring analysis is used as the expected output of the model, and the data are divided randomly in a single pass (75% of the data as the training set and 25% of the data as the test set), and the systemic financial risk early warning models of random forest and BP neural network are constructed in JupyterNotebook, which supports Python language, respectively.

Results of static warning accuracy and recall for random forest model, BP neural network model.

<table>
<thead>
<tr>
<th>Model</th>
<th>Recall rate</th>
<th>Accuracy rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random forest</td>
<td>0.89</td>
<td>0.91</td>
</tr>
<tr>
<td>BP neural network</td>
<td>0.95</td>
<td>0.97</td>
</tr>
</tbody>
</table>

Through the results in Tables 3, it can be concluded that the early warning effect of the deep learning model of BP neural network is better than that of the random forest model, and the accuracy and recall rates of the two types of models on the training set are 0.91 and 0.97, 0.95
and 0.89, respectively. The bp neural network is better than the random forest model with an AUC value of 0.965932. The BP neural network model, whose ROC curve is closer to the upper left corner of the coordinate system than the model, has an AUC value of 0.986235.

Table 4: ROC curves of BP neural network

Table 5: ROC curves for random forests

5 CONCLUSION

This paper employs random forest and BP neural network deep learning models for systemic financial risk early warning in China, comparing and analyzing the early warning effects of different models, as well as comparing the differences in outcome estimation and prediction between traditional econometric models and machine learning and deep learning algorithm models, and identifying the main causes that push up the level of systemic financial risk in China.

The study concludes that, first, in terms of various performance measures and prediction results, machine learning and deep learning algorithm results of ROC curve model evaluation and cross-validation show that machine learning and deep learning algorithms significantly outperform traditional econometric models. The results of ROC curve model evaluation and cross-validation show that machine learning and deep learning algorithm
models perform better than traditional econometric models. In the future, in terms of systemic financial risk early warning, we will continue to optimize and improve the early warning of systemic financial risk in China.

In the future, in terms of systemic financial risk early warning, we will continue to optimize and improve the early warning model of China's systemic financial risk, and introduce more cutting-edge big data analysis technologies such as transfer learning, Meta Learning, Explainable AI and other machine learning and artificial intelligence. To continuously improve the efficiency and performance of systematic In addition, we will revise the data and methods according to the actual needs, and build early warning models that can predict systemic financial risks over a longer period of time on the basis of guaranteeing the accuracy and credibility of prediction.

REFERENCES

Development Potential Analysis and Countermeasure Research of Prepared Dishes Industry Based on SWOT

Xiaoning Jiang
xj2157@nyu.edu
Steinhardt School of Culture, Education, and Human Development, New York University, 82 Washington Square E New York, NY 10003, the United States

Abstract: With the changing public consumption concept and the accelerated pace of life, as well as the changes in the restaurant industry, prepared dishes have developed rapidly in recent years. In order to promote the healthy and sustainable development of the prepared dishes industry, this paper uses SWOT to analyze the advantages, disadvantages, opportunities and threats of its development. Based on the analysis, some countermeasures are formed: Increase the construction of raw material base, enhance the level of scientific and technological, guarantee food quality and safety, improve the construction of cold chain logistics, develop industry standards, guide standardized production, strengthen brand promotion, extend the industrial chain, enhance the value chain. Establishing a whole industry chain digital regulatory system including materials, processing, logistics and distribution is necessary to effectively reduce potential safety risks and improve core competitiveness of prepared dishes enterprises.

Keywords: Prepared Dishes Industry, SWOT, Internal Conditions, External Environment.

1 INTRODUCTION

With the improvement of residents' income, the change of consumption concept and the acceleration of the pace of life, especially in recent years by the impact of the COVID-19, the 'prepared dishes' that can quickly prepare a meal are gaining popularity among people who live and work at home. The market demand is surging.

Prepared dishes refer to dishes prepared with one or more edible agricultural products and their products as raw materials, with or without seasoning and other ingredients, processed by relevant processes, and in certain temperature conditions for storage, transportation and sales, ready to eat or non-ready-to-eat prepackaged dishes. According to different consumption forms, it can be divided into instant food, instant heat, instant cooking and instant preparing [9].

①Instant food: can be eaten directly after opening; ②Instant heat: Food that can be eaten only by heating; ③Instant cooking: The raw food ingredients and necessary condiments that are kept in portions refrigerated or at room temperature, can be immediately into the pot; ④Instant preparing: Raw material food of semi-finished dishes prepared by preliminary processing such as cleaning and cutting.

At present, there are many brands of prepared dishes in the market, and the market demand is growing. It is expected that prepared dishes market in China will maintain a high growth rate of
about 20% in the future [6]. However, as a new industry, there will be obstacles in the process of development. In order to grasp the current problems faced by the development of the prepared dishes industry and explore countermeasures, this paper uses the SWOT to analyze it. Two internal factors S (Strengths) and W (Weaknesses) and two external factors O (Opportunities) and T (Threats) are considered, and two development strategies (SO and WO strategies) and two development countermeasures (ST and WT countermeasures) are formed through the effective matching of the four factors [1].

2 INTERNAL CONDITIONS FOR THE DEVELOPMENT OF THE PREPARED DISHES INDUSTRY

The prepared dishes industry is promising, but there are also some disadvantages hindering its development. Clarify the advantages, disadvantages of the industry and the direction of development, in order to better exploit the strengths and avoid weaknesses, and promote the healthy development of the prepared dishes industry.

2.1 Analysis of Advantageous Factors

2.1.1 Convenient, Time-Saving and Labour Saving

With the accelerated pace of life, most people do not have enough time and energy to cook. Ready-to-heat or ready-to-cook prepared dishes have become their first choice. The prepared dishes has shortened the cooking time and greatly saved time costs.

2.1.2 Reducing Costs and Standardizing the Production

For restaurant operators, prepared dishes not only achieve high efficiency, save labor, time, and kitchen space, but also realize standardized production [10]. It is beneficial for reducing cost.

2.1.3 Raw Materials Supplied Adequately

Skywatch Big Data Report states (2021) that China is a large agricultural producer, and the production of vegetables, meat and other raw materials required for the upstream production of the prepared dishes industry chain is stable and in sufficient supply. The forecast value of raw material output of prepared dishes from 2016 to 2021 is shown in Table 1.

<table>
<thead>
<tr>
<th>Year</th>
<th>Grain (10,000 tons)</th>
<th>Vegetable planting area (10,000 hectares)</th>
<th>Vegetable production (100 million tons)</th>
<th>Meat production (10,000 tons)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2016</td>
<td>66044</td>
<td>1955.31</td>
<td>6.74</td>
<td>8623.33</td>
</tr>
<tr>
<td>2017</td>
<td>66161</td>
<td>1988.11</td>
<td>6.92</td>
<td>8654.43</td>
</tr>
<tr>
<td>2018</td>
<td>65789</td>
<td>2043.89</td>
<td>7.03</td>
<td>8624.63</td>
</tr>
<tr>
<td>2019</td>
<td>66384</td>
<td>2086.27</td>
<td>7.21</td>
<td>7758.78</td>
</tr>
<tr>
<td>2020</td>
<td>66949</td>
<td>2130.56</td>
<td>7.49</td>
<td>7639</td>
</tr>
<tr>
<td>2021</td>
<td>---</td>
<td>2174.43</td>
<td>7.82</td>
<td>7821.84</td>
</tr>
</tbody>
</table>

Data Sources: 2021-2027 China Grain Industry Production and Sales Analysis and Future Prospects Report.
2.2 Analysis of Disadvantageous Factors

2.2.1 The Degree of Deliciousness and Nutrition to be Enhanced

Prepared dishes’ taste and nutrition currently on the market are to be enhanced. Firstly, the traditional high-temperature sterilization are used to ensure product safety, which will lead to nutrient loss and poor taste; Secondly, the prepared dishes mostly use meat as raw materials, low salt, low fat and high dietary fiber products are rare, which will result in unbalanced nutrition.

2.2.2 The Level of Technology and Equipment to be Improved

As a new industry, prepared dishes industry has weak foundation, lacking support of technology and equipment. New packaging materials and technologies suitable for prepared dishes need to be researched and developed [4].

2.2.3 Low Product Taste Reduction

Most of the products have a large difference in taste from the ready-made dishes. It is difficult to keep the ready-made taste for prepared dishes, which results in low consumer repurchase rate and poor market response [8].

2.2.4 Lack of Raw Material Quality Control and Traceability System

Prepared dishes are prone to the problem of judging the quality of ingredients difficulty. Advanced technology and detection means need to be used to identify the material true or false. Therefore, it is necessary to establish a full quality control system as well as a traceability system [4].

3 EXTERNAL ENVIRONMENT OF THE DEVELOPMENT OF THE PREPARED DISHES INDUSTRY

3.1 Analysis of Opportunity Factors

3.1.1 Supported by National Policy Strongly

In recent years, in order to promote the development of the prepared dishes industry, China has promulgated a number of policies related to supporting, encouraging and regulating the prepared dishes industry, such as the Guidance on Accelerating the Cultivation and Development of the Whole Agricultural Industry Chain, issued by the Ministry of Agriculture and Rural. It is pointed out that: innovative development of direct supply from origin of materials, prepared dishes industry, restaurant and take-away, cold chain distribution are need to be encouraged. Also the development and promotion of "raw material base + central kitchen + logistics and distribution", "central kitchen + restaurant" and other models are encouraged [3].
3.1.2 Gradually Increasing Market Demand

![Figure 1: Prepared dishes Market Size and Forecast from 2019 to 2026.](image)


Influenced by the "lazy economy" and the development of the "home-based economy" due to the epidemic, the proportion of prepared dishes in family meals (C-side) are increasing. During the major e-commerce festivals and the Chinese New Year, the prepared dishes products' sales increased with multiples of turnover growth. With the rise of restaurant and take-away industry, the penetration of prepared dishes in restaurants (B-side) has accelerated, and the demand has gradually developed in the direction of "simplification" of dishes, "standardization" of restaurant production, and "acceleration" of take-away orders [7]. The industry expansion is expected to accelerate under the resonance of B+C side demand. According to a survey, the size of China’s prepared dishes industry was 345.9 billion yuan in 2021, up 19.8% year-on-year, and is expected to reach 1072 billion in 2026 (Figure 1).

3.1.3 Contentiously Developing Cold Chain Logistics

Logistics has become more competitive after rapid development, fresh logistics has become a new growth point, and the cost of fresh logistics will be further reduced in the future, providing a more favorable environment for the Internet sales of prepared dishes [6].

3.2 Analysis of Threat Factors

3.2.1 Enterprises Operational Risks

Consumers' demand for abundant prepared dishes products makes it necessary for enterprises to establish multiple production lines, and each line requires multiple links such as new product development, production, transportation and sales, each of which requires the funding. Once the capital chain is broken, the operation of enterprises will be seriously impacted.
3.2.2 Serious Product Homogenization, Fierce Market Competition, Low Profits

Figure 2: Number of sellers & Number of sellers with transactions & Percentage of sellers with transactions.
Data Sources: Prepared dishes industry research report released by the China Federation of Foodstuffs Supply Chain Branch.

The prepared dishes industry is in the early stage of development at present, the permission requirement of manufacturing prepared dishes is less. There are many small and medium-sized enterprises producing similar products. From April 2021 to April 2022, the number of enterprises increases almost threefold (Figure 2). Homogenization is serious and lacking of innovation is prominent. Consumers have many choices when they purchase the same product in the market. So the price is considered as the main factor by the consumers which leads to vicious competition and the overall low interest rate of the industry.

3.2.3 Lack of Relevant Standards and Supervision, Food Safety is Difficult to Guarantee

On the one hand, the nutrients in the prepared dishes are susceptible to microbial infection, which causes spoilage and loss of edible quality. On the other hand, the packaging materials induce component migration in the oil and salt environment of the dishes, which will cause food safety hazards. Corresponding standards and regulations are necessary to be issued for the industry.

3.2.4 Development Policies to be Improved

Prepared dishes industry involves agriculture, processing industry, distribution services and other industries, the development for prepared dishes industry is hindered by the policies corresponding to the industrial ban directory, land, environmental protection, etc. Relevant development policies need to be issued and improved.
3.3 Digital Analysis of the Whole Industrial Chain of Prepared Dishes Enterprises

Figure 3: Registrations of prepared dishes related enterprises & Growth Trend from 2012 to 2021. Data Sources: Big data report "How to make New Year's Eve in Place full of flavor" by Sky Eye Search.

Figure 4: Top 10 enterprises related to prepared dishes in China number of enterprises. Data Sources: China Prepared Dishes Industry Development Trend Research Report 2022 released by iiMedia Research.

In the past 10 years, the number of registered enterprises related to prepared dishes has been reached 69,000 by 2022. The industry is developing rapidly (Figure 3). From the analysis of regional distribution, the number of relevant enterprises in Shandong ranks the first of top ten provinces, with a total of 8200 accounting for 17.75%. The competition is fierce (Figure 4). The period from the purchase materials, processing, transportation to sales of prepared dishes is relatively long. It involves the processing of freezing and refrigeration, there are lot of risks in food safety. The use of digital tool to establish a whole industry chain regulatory system can effectively reduce potential safety risks.

Through data analysis of each links from farm, slaughterhouse, food processing manufacturer, cold-chain logistics, distribution and sales, differentiated competition advantage in prepared
dishes enterprises will be achieved. Incorporating these scene data, behavioral data, transaction data, and all original data into a big real-time monitoring and management data analysis system, integrating the upstream and downstream information of the industry, building a whole-chain prepared dishes industry digital system including processing, safety supervision and industrial services, a wide and practical digital product matrix of prepared dishes will be formed. The system covers digital integration capability for upstream enterprises (vegetable planting, poultry farming, aquaculture, spice industry, etc.), digital supply chain capability, for midstream enterprises (digital procurement, storage, sorting, logistics, settlement, production, central kitchen, traceability), digital marketing and sales capability for downstream enterprises (CRM marketing, community e-commerce, third-party e-commerce: TikTok, Taobao, Pinduoduo, etc.), digital decision-making capability. Based on the above, business scale and core competitiveness of prepared dishes enterprises will be improved.

4 CONCLUSION AND COUNTERMEASURES

Based on the SWOT analysis, various strategies and countermeasures for the development of the prepared dishes industry are summarized as follows:

4.1 Enhance the Construction of Raw Material Base

In good industrial town, modern agricultural industrial park, guide the prepared dishes processing enterprises to establish a number of standardized green production base. Explore “farmers + production base + catering enterprises” and “farmers + production base + electric business” and other modes to realize the industrial integration development and help farmers to increase income.

4.2 Enhance the Level of Technological Support

Scientific research institutes and promotion institutions were organized to conduct research on the key processing and bottleneck constraints of prepared dishes industry, key technical research was carried out, and a number of new practical technologies such as green packaging, rapid pre-cooling, energy-saving drying, which integrated automatic and precise control, were developed to provide technical support for the production of prepared dishes.

4.3 Guarantee Food Quality and Safety

Government support prepared dishes processing enterprises to participate in quality management and food safety control system certification, and enhance the quality control ability. Encourage prepared dishes enterprises to build food safety management system which can realize entire traceability, and promote the systematization, institutionalization and standardization of prepared dishes processing.

4.4 Improve Cold Chain Logistics

In the vegetable dominant production areas, vegetable industry clusters, key vegetable production base, support building cold storage and pre-cooling facilities and to improve the quality of raw material for prepared dishes. Relying on prepared dishes processing enterprises,
actively cultivate cold-chain logistics, packaging, distribution and other professional service entities to improve the commercialization processing capacity such as pre-cooling.

4.5 Develop Industry Standards and Guide Standardized Production

Government encourage market supervision departments and industry associations to co-establish standard, guide standardized production to develop prepared dishes healthily and promote standardized production vigorously. Meanwhile, encourage and guide enterprises to manufacture abiding by advanced standards which is coordinated with the upstream agricultural standard system to ensure food safety and regulate industrial development.

4.6 Strengthen the Promotion of the Brand

Around scientific and technological innovation, waste reduction, green development, The typical enterprise and product brand were selected and promoted. Fully exploit the typical model and successful practices, conduct a series of publicity reports, and form a good atmosphere for the development of prepared dishes industry.

Table2: prepared dishes industry development strategy SWOT analysis matrix.

<table>
<thead>
<tr>
<th>External environmental factors</th>
<th>Opportunity Factor (O)</th>
<th>Threat Factors (T)</th>
</tr>
</thead>
<tbody>
<tr>
<td>① Supported by National policy strongly. ② Gradually increasing market demand. ③ Contentiously developing cold chain logistics .</td>
<td>① Enterprises operational risks. ② Serious product homogenization, fierce market competition, low profits. ③ Lack of relevant standards and supervision, food safety is difficult to guarantee. ④ Development policies to be improved.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Internal condition factors</th>
<th>Advantageous factors (S)</th>
<th>Disadvantageous factors (W)</th>
</tr>
</thead>
<tbody>
<tr>
<td>① Convenient, time-saving and labour-saving. ② Reducing costs and standardizing the production. ③ Raw materials supplied adequately.</td>
<td>(Take advantages and seize the opportunity) ① Extend the industrial chain and guide vegetable processing enterprises to upgrade to prepared dishes manufacturing enterprises. ② Promote the value chain and ensure the high-quality development of the industry.</td>
<td>(Seize the opportunity, change the disadvantage) ① Develop industry standards and guide standardized production. ② Develop supportive policies and improve the development environment. ③ Enhance the construction of raw material base.</td>
</tr>
<tr>
<td>① Enterprising operational risks. ② Serious product homogenization, fierce market competition, low profits. ③ Lack of relevant standards and supervision, food safety is difficult to guarantee. ④ Development policies to be improved.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SO Strategy: Proactive Attack</td>
<td>WT Countermeasures: Defend or Retreat</td>
<td></td>
</tr>
<tr>
<td>ST Strategy: Corresponding Defensive Style</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WO Countermeasures: Progressive</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Exploit advantages, avoid threats) ① Develop industry standards and guide standardized production. ② Develop supportive policies and improve the development environment. ③ Enhance the construction of raw material base.
① The degree of deliciousness and nutrition to be enhanced.
② The level of technology and equipment to be improved.
③ Low product taste reduction.
④ Lack of raw material quality control and traceability system.

① Improve cold chain logistics.
② Strengthen the promotion of the brand.
① Enhance the level of technological support.
② Guarantee food quality and safety.

REFERENCES

Analysis of the Relationship Between ETF Volatility and Liquidity Based on ARMA-GARCH Model

Qingyuan Feng
Email: hwhqy@163.com
Canvard College, Beijing Technology and Business University, Beijing, 101118, China

Abstract: Based on ARMA-GARCH model, this paper takes Amihud's non liquidity ratio as an indicator to measure ETF liquidity, and makes an empirical analysis on the impact of ETF liquidity on return volatility in China. The analysis results show that the weak liquidity of ETFs has a positive impact on the volatility of returns, and it also has a certain explanatory effect on the risk premium of ETFs. However, the explanatory power of liquidity is limited, and there are other factors that affect the volatility of ETFs' returns. Finally, some policy suggestions are given.

Keywords: Volatility, ETF, ARMA-GARCH Model.

1 INTRODUCTION

Exchange Traded Fund (ETF), as an important asset allocation tool for market investors, has been favored by investors for many years since it was launched in the United States in 1993. In 2004, the first ETF product in China, Huaxia CSI 50ETF, was established. Since then, the ETF market in China has shown a vigorous development trend [7]. According to the Shanghai Stock Exchange ETF Industry Development Report (2022) released by the Shanghai Stock Exchange, by the end of 2021, there were 635 ETFs listed for trading in China, with the total assets reaching 1405.2 billion yuan, and the annual turnover of ETFs in Shanghai Stock Exchange alone exceeded 15 trillion yuan.

However, because ETFs have the characteristics of both stocks and index funds, liquidity problems may arise when trading in the secondary market of the exchange. For ETFs, liquidity is very important. Only better liquidity can attract customers, because their products are highly homogeneous. Even though the overall ETF is growing, in fact, except for the head ETF, the liquidity and trading volume of other parts are not good. The lack of liquidity of ETFs that are unpopular with investors may cause market makers to be limited in developing appropriate markets, and then increase the transaction costs and risks borne by such ETF investors [3]. Therefore, it is significant for investors to avoid unnecessary additional risks by studying the impact of ETF liquidity on its volatility [8].
2 THEORY INTRODUCTION

2.1 Selection of Liquidity Indicators

The Amihud measure is selected as an indicator to measure liquidity. Since the yield data used in this paper are daily data, the Amihud measure is calculated according to the following formula:

\[ \text{Amihud}_{it} = \frac{|R_{it} - R_{i,t-1}|}{V_{it}} \]

where \( R_{it} \) are the yields of securities \( i \) on trading day \( t \) and \( V_{it} \) are the trading volumes of securities \( i \) on trading day \( t \) [1]. After the ADF test of Amihud measure, we can find that the Amihud index series of ETF funds from January 1, 2015 to December 31, 2021 is stable.

2.2 ARMA-GARCH Model Introduction

ARCH model is also called "autoregressive conditional heteroscedasticity model". This model is proposed by Engle to solve the conditional heteroscedasticity of data and applied to the study of volatility. However, the lag term of the variance equation is sometimes large, and the ideal volatility equation cannot be obtained by using the ARCH model. Therefore, in 1986, Bollerslev proposed the GARCH (p, q) model to improve the defects of the ARCH model. Because most of the financial time series data have the phenomenon of volatility aggregation, showing the characteristics of thick tail distribution, the use of this model can effectively eliminate the excessive peak problem caused by data [2]. GARCH (1,1) model is often used as a modeling tool in practical problems [2]. This paper adopts the widely used GARCH (1,1) model, and adds the weak liquidity index represented by Amihud measure to the conditional variance equation of GARCH (1,1) model, and removes the impact of trading volume by extracting the residual of Amihud measure on the linear regression of trading volume. Thus, GARCH (1,1) model with weak liquidity index is obtained:

\[ r_t = c + \text{ARMA}(p,q) + \epsilon_t \]  
\[ \sigma^2_t = \alpha_0 + \alpha_1 \sigma^2_{t-1} + \beta \epsilon_{t-1} \]  
\[ \text{Amihud}_{t} = \delta_0 + \delta_1 V_{t} + \epsilon_t \]

Wherein, equation (1) is called mean value equation and equation (2) is conditional variance equation. \( \epsilon_t \) follows an unknown distribution with mean value of 0 and variance of \( \sigma^2_t \). \( \sigma^2_t \) represents the volatility of the yield, namely the conditional variance. The model has constraints \( 0 < \alpha_1 + \alpha_2 < 1 \), which is used to ensure that the unconditional variance of \( \epsilon \) satisfying the model is limited and unchanged, while the conditional variance \( \sigma^2_t \) can change over time [4].
3 EMPIRICAL ANALYSIS

3.1 Basic Statistical Characteristics of ETF Fund Yield

The research object of this paper is China's Exchange Traded Fund (ETFs). The data use the data set from January 1, 2015 to December 31, 2021 in Guotai An CSMAR database. Excluding the missing data of holidays or individual dates, the daily returns of 1303 ETF funds considering cash dividends in 1705 trading days were obtained.

After cross sectional average of the returns of these 1303 ETFs, we can get a series of returns. Figure 1 shows the general trend of the ETF return series in 1705 trading days. From the time series diagram of ETF yield, it can be seen intuitively that the large and small fluctuations of ETF yield tend to gather in different periods. Based on this, it can be preliminarily inferred that the volatility of ETF yield has aggregation, and because the amplitude of volatility is inconsistent, it can be judged that there may be heteroscedasticity in the yield series.

Figure 1 ETF Fund Yield Time Series

Descriptive statistics of ETF yield series are shown in Table 1 and Figure 2. It can be seen from Table 1 that the mean and standard deviation of ETF returns are small, the skewness is negative, and the kurtosis is far greater than 3, indicating that this distribution has a long left tail and a fatter tail compared with the normal distribution. The Jarque Brea statistic is 5416.71, which is significant at the 1% level. This data also shows that the ETF yield series does not obey the normal distribution. Figure 2 is a Q-Q scatter chart of the yield series. It can be found that the tail of the yield series deviates from the diagonal seriously, which indicates that the distribution of ETF returns has the characteristics of a thick tail, which again indicates that the yield series does not obey the normal distribution. The unit root test of the yield series shows that the t statistic obtained is significant at the 1% level, indicating that the yield series is stable [9].

Table 1 Descriptive Statistics of ETF yield

<table>
<thead>
<tr>
<th>Mean</th>
<th>Standard Deviations</th>
<th>skewness</th>
<th>kurtosis</th>
<th>J-B</th>
<th>ADF</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00062</td>
<td>0.011258</td>
<td>-0.311215</td>
<td>11.70974</td>
<td>5416.710**</td>
<td>-34.29044**</td>
</tr>
</tbody>
</table>
Next, we use the autocorrelation and partial autocorrelation coefficient graph of the yield series and Ljung Box Q statistics to determine whether the series has serial correlation. Figure 3 shows the autocorrelation and partial autocorrelation coefficients of ETF yield series. It can be seen from the figure that the autocorrelation and partial autocorrelation values of each order lag are very small, and the corresponding P value is also less than 0.001, so it is judged that there is significant autocorrelation in the yield series [10].

![Figure 2 Q-Q Scatter Chart of ETF Yield](image)

To sum up, the ETF yield series shows these statistical characteristics: peak and fat tail, autocorrelation and bias. For the autocorrelation of the yield series, this paper plans to introduce the conditional mean model - ARMA model to eliminate the series correlation. However, the index GARCH (EGARCH) model proposed by Nelson (1991), the asymmetric power ARCH (APARCH) model proposed by Ding, Granger and Engle (1993) all allow positive and negative asset returns to have an asymmetric impact on volatility, which can show better results in biased description and analysis [5].

### 3.2 Construction of mean value equation by ARMA model

First, we need to determine the form of the mean value equation. According to the autocorrelation and partial autocorrelation coefficient graph of the yield (Figure 3), we can roughly see the maximum values of the two parameters p and q in the ARMA (p, q) model. Taking max (p)=5 and max (q)=5, after comparing AIC, SC and HQC values of 36 models, the
optimal model was determined to be ARMA (4,3) using the principle of minimizing information criteria. Table 2 shows the regression results of ARMA (4,3) model.

<table>
<thead>
<tr>
<th>Term</th>
<th>Coefficient</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(1)</td>
<td>-0.599912</td>
<td>-27.12507</td>
<td>0.0000</td>
</tr>
<tr>
<td>AR(2)</td>
<td>-0.614815</td>
<td>-25.51083</td>
<td>0.0000</td>
</tr>
<tr>
<td>AR(3)</td>
<td>-0.746230</td>
<td>-41.04747</td>
<td>0.0000</td>
</tr>
<tr>
<td>AR(4)</td>
<td>0.213533</td>
<td>16.05401</td>
<td>0.0000</td>
</tr>
<tr>
<td>MA(1)</td>
<td>0.819291</td>
<td>42.15887</td>
<td>0.0000</td>
</tr>
<tr>
<td>MA(2)</td>
<td>0.741298</td>
<td>28.70628</td>
<td>0.0000</td>
</tr>
<tr>
<td>MA(3)</td>
<td>0.909103</td>
<td>51.90958</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

The mean value equation established by ARMA (4,3) is given by the following equation:

\[ r_t = -0.600r_{t-1} - 0.615r_{t-2} - 0.746r_{t-3} + 0.214r_{t-4} + u_t + 0.819u_{t-1} + 0.741u_{t-2} + 0.909u_{t-3} \]

Figure 4 shows the residual sequence diagram of the mean value equation. It can be preliminarily judged from the observation chart that there may be conditional heteroscedasticity effect in the residual sequence.

Next, test whether the residual of the equation has the ARCH effect according to the Lagrange multiplier (LM). By using the information criterion minimization principle, the lag order is determined to be 5, and the ARCH-LM test results are shown in Table 3.

<table>
<thead>
<tr>
<th>LM</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>408.8875</td>
<td>0.0000</td>
</tr>
</tbody>
</table>
The test results show that the LM statistic is very significant at the 1% level, that is, the residual sequence has an obvious ARCH effect.

### 3.3 Establishment of benchmark GARCH (1,1) model

When constructing the GARCH (1,1) model of benchmark income, as the tail distribution is unknown, first compare the GARCH (1,1) models under different tail distribution assumptions to determine the optimal model. Here, three different tail distributions are compared, namely, normal distribution, student-t distribution and generalized error distribution (GED). Table 4 lists the AIC, SC and HQC values of the model under the three tail distributions.

<table>
<thead>
<tr>
<th>Tail distribution</th>
<th>AIC</th>
<th>SC</th>
<th>HQC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>-6.597068</td>
<td>-6.587494</td>
<td>-6.593524</td>
</tr>
<tr>
<td>Student-t</td>
<td>-6.672364</td>
<td>-6.659599</td>
<td>-6.667639</td>
</tr>
<tr>
<td>GED</td>
<td>-6.667961</td>
<td>-6.655195</td>
<td>-6.663236</td>
</tr>
</tbody>
</table>

It is not difficult to see that GARCH (1,1) model based on Student-t distribution is optimal.

### Table 5 Regression Results of GARCH Benchmark(1,1) Model

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Z</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_0$</td>
<td>8.45E-07**</td>
<td>2.553372</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>0.079742***</td>
<td>5.921469</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>0.915373***</td>
<td>74.21357</td>
</tr>
<tr>
<td>$\alpha_1 + \alpha_2$</td>
<td>0.995115</td>
<td></td>
</tr>
</tbody>
</table>

Table 5 shows the estimation results of the GARCH (1,1) model of benchmark returns when the tail distribution follows the t distribution. It can be seen that in the benchmark model, $\alpha_1$ and $\alpha_2$ are significant at 1% or higher significance level. The sum of $\alpha_1$ and $\alpha_2$ is highly close to 1, which indicates that the return sequence of ETF has a high sustainability feature; $\alpha_2$ is greater than 0.9 and highly significant, indicating that the market memory is strong and the impact of conditional variance is more lasting.

In order to test whether GARCH (1,1) model eliminates ARCH effect, this paper conducts autocorrelation test again, this time it is the residual sequence of the model. The following table shows the ARCH-LM test results of the residual sequence, as shown in Table 6.

### Table 6 ARCH-LM Results of GARCH (1,1) residual

<table>
<thead>
<tr>
<th>LM</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.189669</td>
<td>0.2069</td>
</tr>
</tbody>
</table>

The test results show that the model effectively eliminates the ARCH effect of residual sequence.
3.4 GARCH (1,1) model with liquidity indicators

In order to study the impact of liquidity on the volatility of ETF returns, we will add liquidity indicators to the benchmark GARCH (1,1) model of ETF returns. Amihud measure is selected as the liquidity indicator here to measure the illiquidity of securities [6].

First, the stability of Amihud measure and trading volume series is tested. The results are shown in Table 7. At the 5% significance level, it can be considered that the two sequences are stable and single integer of the same order.

<table>
<thead>
<tr>
<th></th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amihud</td>
<td>-7.034456***</td>
<td>0.0000</td>
</tr>
<tr>
<td>Transaction volume</td>
<td>-3.843325***</td>
<td>0.0026</td>
</tr>
</tbody>
</table>

Use the Engle Granger two-step method to carry out cointegration test for two variable sequences. The OLS is used to regress the two sequences, extract the residual sequence, and test. Here, ADF test is selected. The results show that the T value is -12.0519, and the significance level is 1%, which is significantly stable, the residual sequence is stable, and there is cointegration between Amihud measure and trading volume.

We carry out OLS regression in equation (3) and extract residual sequence. The estimated results of parameters of GARCH (1,1) model including liquidity indicators obtained by adding residual series to GARCH (1,1) model are shown in Table (8).

The results show that the coefficient of weak current term β is positive and significant at the 5% significance level; After adding weak liquidity indicators, α₁ and α₂ in the model is still highly significant, but its corresponding z values have decreased, and the sum of α₁ and α₂ used to reflect the persistence of volatility has also decreased to a certain extent. The value of α₂ is still large and significant, which indicates that the aggregation and persistence of volatility are still very obvious, and there are still unknown factors that can be used to explain the volatility of ETF fund returns.

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>Z</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>α₀</td>
<td>1.11E-06***</td>
<td>3.026290</td>
<td>0.0025</td>
</tr>
<tr>
<td>α₁</td>
<td>0.075679***</td>
<td>5.667164</td>
<td>0.0000</td>
</tr>
<tr>
<td>α₂</td>
<td>0.915487***</td>
<td>73.43036</td>
<td>0.0000</td>
</tr>
<tr>
<td>β</td>
<td>0.061765**</td>
<td>2.193103</td>
<td>0.0283</td>
</tr>
<tr>
<td>α₁ + α₂</td>
<td>0.991166</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To sum up, the liquidity of ETF can be used to explain the volatility of its yield. Weak liquidity will increase the volatility of ETF's yield, while when liquidity increases, the volatility of ETF's yield will weaken; However, liquidity has limited ability to explain ETF yield volatility. The introduction of weak liquidity indicators cannot significantly reduce the
persistence and aggregation of ETF yield volatility. There are still other factors behind it that can explain the volatility of ETF yield; There is an obvious risk premium in China's ETFs, and due to the leverage effect, when the ETFs have insufficient liquidity, it is more likely to cause large fluctuations in their yield, which will lead to investors taking greater risks and expanding transaction costs.

4 CONCLUSIONS AND RECOMMENDATIONS

The following conclusions are drawn from the analysis of modeling results:

(1) The volatility of China's ETF returns has obvious characteristics of aggregation and persistence, that is, larger volatility tends to be accompanied by larger volatility, smaller volatility tends to be accompanied by smaller volatility, and periods of high volatility and low volatility will alternate;

(2) Liquidity is an important factor to explain ETF yield volatility. Weak liquidity has a positive impact on ETF yield volatility. Weakening ETF liquidity will lead to an increase in ETF yield volatility. Accordingly, when liquidity is enhanced, ETF yield volatility will weaken;

(3) There is leverage effect in the volatility of ETF yield in China, and the negative impact on the volatility is often greater than the positive impact on the volatility. Therefore, when the liquidity is weakened, it is more likely to cause a large range of yield volatility, making investors bear greater risks;

(4) ETFs in China have an obvious risk premium, and the liquidity of ETFs is one of the factors that explain the risk premium. Therefore, if the ETF has insufficient liquidity, the weakening of liquidity will lead to the increase of yield volatility and risk, which will lead to the expansion of transaction costs.

(5) Liquidity has limited ability to explain ETF yield volatility, and the aggregation and persistence of volatility cannot be fully explained by liquidity. There are still unknown other explanatory factors behind it that can be used to explain ETF yield volatility;

Based on the research results, relevant policy recommendations are proposed:

(1) In order to further improve the liquidity of ETFs, some fund companies concerned are expected to do something, mainly in a series of practices such as increasing market makers or adjusting the minimum subscription and redemption units. So as to improve the trading volume of ETF and expand the scale of ETF, so as to promote the better and faster development of ETF;

(2) Due to the head effect and first mover advantage of China's ETF market, in the context of the increase in the size of China's ETF market in recent years and the cooling of new funds, focusing on liquidity will help build investor confidence and the recovery of ETF market issuance; For funds with seriously insufficient liquidity, relevant regulations should be formulated to clear them from the market in a timely manner.
REFERENCES

The Gratification Driving User Attitude and Continuance Use of Mobile Payment Services in China

Liming Zhang¹,a, Ruochen Hu¹, Kuankuan Luo¹,b*
982578492@qq.com,a, luokuankuan@scu.edu.cn*
Business School, Sichuan University, Chengdu, People’s Republic of China¹

Abstract: Purpose/Meaning: This study aims to investigate the impact of gratification on users’ attitudes and continuance use of mobile payment services in China. Process/Method: Data was collected from conveniently sampled 200 users of mobile payment services in China. A questionnaire, which mainly contains five-point Likert scale questions, was used to collect the data. The study adopted the Uses and Gratification (U&G) theory and was analyzed with the Structural Equation Modelling. Result/Conclusion: The findings show that cognition gratification, convenience gratification and usefulness gratification can significantly impact users’ attitudes towards mobile payment services. In addition, user attitudes can also significantly influence the continuance use intention of mobile payment services. Keywords: Mobile payment, Gratification, User attitude, Continuance use, Technology acceptance behavior.

1 INTRODUCTION

Due to the ubiquitous characteristic of mobile devices, a variety of new financial services have created [6, 13]. For fintech services in developing economies, the mobile payment service market is often the fastest growing segment due to its convenient payment characteristics [15]. Mobile payment uses the mobile communication terminals and device to conduct bank transfer, payment, shopping and other transaction activities through the SMS (short message service), IVR (the interactive voice), WAP (wireless application protocol) and many other communication methods. As a new payment method, mobile payment has the advantages of convenience, security and reliability.

According to Chinese Ministry of Information Industry, there were 547 million mobile phone users in China in 2007, ranking first in the world. In China, the application scenarios of mobile payment are also constantly enriched, from daily bill payment to banking business, from public transportation to wealth management and shopping, all can be completed through mobile payment. Payment forms are also increasingly diversified, whether it is fingerprint payment or face-scanning payment, which makes consumption more convenient.

Although mobile payment services have become a hot topic in academia, there is little research on whether and how users’ usage behaviors change after they first adopt mobile payments. [6, 25]. To fill the gaps in the previous literature, this study chose to adopt the use and
gratification theory (U&G)\footnote{14} to identify various key factors that improve users’ attitudes towards mobile payments and drive continuance use in China.

## 2 LITERATURE REVIEW

### 2.1 Use and Gratification Theory

Use and Gratification Theory is a way used to help understand how individual adopt and use technology\footnote{11}, which is proposed on the basis of technology acceptance model, computer use model, incentive model, rational behavior and planned behavior theory. U&G theory states that individuals will obtain or seek gratification when they use their preferred media channels. The gratification is generally recognized as the potential benefits of using the medium when individuals choose a channel\footnote{20, 21}. When users are satisfied after using a particular medium, they have a positive attitude towards the medium and vice versa, which affects their willingness to choose that medium again in the future\footnote{18}. The literature suggests that user gratification categories include gratification from information acquisition and understanding of the surrounding environment (Cognition), gratification from pleasant experiences (Hedonic), gratification sought from enhanced trust, personal gratification Trust, connection with friends and family (Integration)\footnote{5, 12}.

#### 2.1.1 Cognitive

Cognitive gratification is defined as “the degree to which people believe that using a particular system can improve their performance”\footnote{8}. Venkatesh et al. (2012) also refer to it as performance expectation. There were already researchers found that cognitive gratification is related to product and event information, curiosity, consultation and knowledge acquisition\footnote{24}. There was a study find that there is a significant correlation between cognitive gratification and mobile payment technology, which is embodied in the positive impact of cognitive value on users’ acceptance of mobile payment\footnote{25}. In addition, Chang et al. (2016) also found that users’ attitudes towards mobile payments can be significantly impacted by utilitarian gratification.

#### 2.1.2 Hedonic

Extant literature refers to hedonic gratification as the gratification of enhanced pleasure experiences. There is a strong correlation between hedonic gratification and stress release needs\footnote{31}. There were also studies found that hedonic gratification is related to the reduction of boredom and pleasure\footnote{10}. According to these studies, Ha et al. (2015) found that hedonic gratification played an important role in impacting Korean users’ attitudes towards using social networking sites. Likewise, Azam (2015) stated that, in Saudi Arabia, hedonic gratification can directly influence students’ attitudes towards website use. However, Ozturk et al. (2017) argued that mobile payments will only bring functional benefits to users, which is the reason that different studies disagree on hedonic gratification.

#### 2.1.3 Convenience

Dewan and Chen (2005) believed that the innovative and high-tech nature of mobile payments may bring convenience to users. According to extant research, convenience has been
unanimously recognized as an important factor impacting users’ attitude towards mobile payment \cite{15, 28}. This view is also supported by a study by Ozturk et al. (2017) on user acceptance of mobile payment technologies. They believe that users who feel the convenience of using mobile payments are more willing to use this technology.

### 2.1.4 Ease of Use

Perceived ease of use refers to the degree to which a person thinks that using a technology can save effort \cite{8}. After being widely implemented and validated, ease of use gratification emerged as an important factor in explaining the use of the technology \cite{3, 16}. Furthermore, Mun et al. (2017) found that ease of use gratification is an important factor influencing mobile payment services for millennials. In 2013, Shaw also conducted a study of the use of mobile wallets by Canadian business school students. However, according to the study setting, ease of use gratification had no positive effect on intention to use a mobile wallet (Shaw, 2014).

### 2.1.5 Usefulness

Subsequent literature has demonstrated that usefulness gratification is an important factor affecting users’ attitudes toward technology use \cite{3, 35}. In China, perceived risk is a key factor for users to judge the usefulness of a technology. Based on high-speed communication technologies, mobile payment has supported smart devices to conduct real-time data collection and comparative analysis of personal biometric characteristics, including face and fingerprint. The trusted computing technologies create has enhanced users’ trust to mobile payment, which also increase the usefulness of mobile payment technologies.

### 2.2 Attitude of Mobile Payment

Ajzen (2001) refers to attitude as a total score of mental state, which captures dimensions such as good and bad, beneficial and harmful, pleasant and unpleasant. There have been studies to explore how attitudes influence user behavior. For example, Phonthanukitthaworn et al. (2015) discussed the impact of utilitarianism on attitudes towards mobile payment usage. Additionally, other studies suggested that actual usage behavior can be affected by the attitude whether it is positive or negative \cite{7, 25}. Although there have been studies which confirmed the influence of attitude at first use on user acceptance in the field of mobile payments \cite{10, 19, 29}. However, there are few studies examining the impact of attitudes on continuance use of mobile payments.

### 3 RESEARCH METHODS AND MATERIALS

The research about "The Gratification Driving User Attitude and Continuance Use of Mobile Payment Services" is a quantitative study using online questionnaires as a data collection tool and analyzed by statistical procedures.

#### 3.1 Research Framework

According to the theoretical framework of previous research, we can construct the conceptual framework of this research through 7 variables, which are Cognitive Gratification (CG), Hedonic Gratification (HG), Convenience Gratification (CO), Ease of Use Gratification (EG),
Usefulness Gratification (UG), Attitude of Mobile Payment (AT) and Continuance Use (CU). And the framework is shown as Figure 1.

![Conceptual framework](image)

**Figure 1: Conceptual framework**

From the conceptual framework, the hypothesis of the variables are:

H1a: Cognitive Gratification is significantly impacting on user attitude of mobile payment in China

H1b: Hedonic Gratification is significantly impacting on user attitude of mobile payment in China

H1c: Convenience Gratification is significantly impacting on user attitude of mobile payment in China

H1d: Ease of Use Gratification is significantly impacting on user attitude of mobile payment in China

H1e: Usefulness Gratification is significantly impacting on user attitude of mobile payment in China

H2: User attitude of mobile payment is significantly impacting on Continuance Use of mobile payment in China

### 3.2. Research Methodology

#### 3.2.1 Sample and Data Collection

In this study, the proposed data collection tool is to conduct a questionnaire survey through the online platform "Questionnaire Star". The questionnaires were distributed to the target population using the non-probability method of convenience sampling technique in this study.

The questionnaire used in this study was derived from a questionnaire adapted from Alhassan et al. (2020). Consists of 3 parts with 35 questions, of which 2 screening multiple-choice questions aim to identify target respondents, 3 demographic multiple-choice questions aim to collect demographic data on mobile payment users in China, and 30 five-point Likert scales. The questions were designed to collect data on factors influencing user gratification, attitudes
towards mobile payments, and continued willingness. The Likert scale consists of 5 scales, 1 is strongly disagree, 2 is disagree, 3 is neutral, 4 is strongly agree, and 5 is strongly agree.

The target population of this study is the total users of mobile payment in China.

A widely used minimum sample size estimation method is the "10-times rule" method [26]. The "10-times rule" emphasizes that the minimum sample size of a study should exceed "10-times" the maximum number of internal or external model links that lead to any structure. Therefore, the minimum sample size required for this study is 50 respondents. A total of 212 questionnaires were distributed in this study, which met the minimum sample size requirements.

### 3.2.2 Method of Data Analysis

In this study, frequency and percentage descriptive statistical methods were used to analyze user demographic characteristics. Structural Equation Modelling was used to examine the significant effects between five gratification variables, Chinese mobile payment user attitudes and their Continuance Use. As a multivariate data analysis technique, SEM is widely adopted by IS research and is useful in studies with limited respondents and skewed data distribution.

### 4 RESULTS AND DISCUSSION

The purpose of this study is to investigate the factors that influence the attitude and continued usage of mobile payment users in China. Data was collected through an online questionnaire, and 200 respondents were finally obtained after deducting respondents who were incomplete and did not meet the screening requirements. The results were as follows:

#### 4.1 Demographic Profile Summary

The demographic profiles of the participants are showed in Table 1. The collected data shows that all respondents have used mobile payment services, among which WeChat Pay (N=158) and Alipay (N=42) are the main ones. In China, the vast majority of users choose to use WeChat Pay or Alipay, which are operated by Tencent and Alibaba respectively. All users with a smartphone can register and use it on the corresponding platform.

<table>
<thead>
<tr>
<th>Demographic</th>
<th>Characteristics</th>
<th>No.</th>
<th>(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>Male</td>
<td>78</td>
<td>39.0%</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>122</td>
<td>61.0%</td>
</tr>
<tr>
<td>Age</td>
<td>Below 18 years</td>
<td>2</td>
<td>1.0%</td>
</tr>
<tr>
<td></td>
<td>18-26 years</td>
<td>107</td>
<td>53.5%</td>
</tr>
<tr>
<td></td>
<td>27-35 years</td>
<td>48</td>
<td>24.0%</td>
</tr>
<tr>
<td></td>
<td>More than 36 years</td>
<td>43</td>
<td>23.5%</td>
</tr>
<tr>
<td>Most Used Mobile Payment Platform</td>
<td>Wechat Pay</td>
<td>152</td>
<td>79.0%</td>
</tr>
<tr>
<td></td>
<td>Ali Pay</td>
<td>42</td>
<td>21.0%</td>
</tr>
</tbody>
</table>

Table 1: Demographic characteristics of respondents
As can be seen in Table 1, taking into account a total of 194 respondents, the Chinese mobile payment users described by the results are 39.0% male (78 respondents) and 61.0% female (122 respondents). 18-26 years old dominated, accounting for 53.5% (107 respondents), followed by 24.0% (48 respondents) between 27-35 years old, 21.5% (43 respondents) over 36 years old, and 1.0% (2 respondents) below 18 years old.

4.2 Reliability Test

This study conducted a pilot test before the formal launch of the questionnaire. In the pilot test, a questionnaire survey was conducted among 30 mobile payment users to investigate the internal shortcomings, validity and reliability of the questionnaire.

Indicator reliability is described as the degree to which "a variable or set of variables is consistent in what it is intended to measure" [33]. A pilot test was performed to check indicator loadings of each variable. However, not all indicators have significant loads (more than 0.7) on their corresponding latent variables. As a result, they are removed from the model. The results of pilot test showed that HG4 and UG1 made a bad influence about the reliability of variables whose indicator loadings were 0.681 and 0.662. This may mean that these questions are not suitable for Chinese mobile payment users. After delete these questions, the results of the model are shown as Figure 2.

![Indicator loading after factor analysis](image)

**Figure 2:** Indicator loading after factor analysis

4.3 Structural Model Assessment

4.3.1 Assessment for Multicollinearity

This study first examined the Multicollinearity. From Table 2, all factors have a VIF value below 5, which indicates that this study has no problem about Multicollinearity.
Table 2: Multicollinearity statistics (inner VIF)

<table>
<thead>
<tr>
<th></th>
<th>AT</th>
<th>CU</th>
<th>CG</th>
<th>HG</th>
<th>CO</th>
<th>EG</th>
<th>UG</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT</td>
<td>1.827</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CU</td>
<td></td>
<td>2.788</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG</td>
<td></td>
<td></td>
<td>1.308</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HG</td>
<td></td>
<td></td>
<td></td>
<td>3.678</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CO</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.465</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.908</td>
<td></td>
</tr>
<tr>
<td>UG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4.358</td>
</tr>
</tbody>
</table>

4.3.2 Assessment for Path Coefficient

To evaluate the significance of the path coefficient, this study run a bootstrapping algorithm. As 5000 subsamples at 0.1 and 95% confidence interval, the findings are shown in Table 3.

Table 3: Direct relationship for hypotheses testing

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Relationship</th>
<th>Std beta</th>
<th>Std error</th>
<th>t-value</th>
<th>Inference</th>
<th>5% CILL</th>
<th>95% CIUL</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1a</td>
<td>CG-&gt;AT</td>
<td>0.371</td>
<td>0.067</td>
<td>4.569</td>
<td>Accepted</td>
<td>0.176</td>
<td>0.378</td>
</tr>
<tr>
<td>H1b</td>
<td>HG-&gt;AT</td>
<td>-0.006</td>
<td>0.09</td>
<td>0.193</td>
<td>Rejected</td>
<td>-0.198</td>
<td>0.103</td>
</tr>
<tr>
<td>H1c</td>
<td>CO-&gt;AT</td>
<td>0.42</td>
<td>0.079</td>
<td>5.817**</td>
<td>Accepted</td>
<td>0.395</td>
<td>0.798</td>
</tr>
<tr>
<td>H1d</td>
<td>EG-&gt;AT</td>
<td>0.073</td>
<td>0.078</td>
<td>0.207</td>
<td>Rejected</td>
<td>-0.033</td>
<td>0.188</td>
</tr>
<tr>
<td>H1e</td>
<td>UG-&gt;AT</td>
<td>0.124</td>
<td>0.042</td>
<td>2.84**</td>
<td>Accepted</td>
<td>0.04</td>
<td>0.23</td>
</tr>
<tr>
<td>H2</td>
<td>AT-&gt;UG</td>
<td>0.819</td>
<td>0.041</td>
<td>19.389**</td>
<td>Accepted</td>
<td>0.696</td>
<td>0.767</td>
</tr>
</tbody>
</table>

4.3.3 Assessment for Goodness of Fit

To assess the goodness of fit, this study choose the GOF test. In the test, R^2 determination coefficient (R^2) can measures the explanatory power of the model, which makes it a widely used criteria. In IS research, R^2 values of 0.670 and above can be considers as substantial respectively. So the finding in Table 4 shows that the R^2 of the model is substantial.

Table 4: R-squared

<table>
<thead>
<tr>
<th>Dependent constructs</th>
<th>R^2</th>
<th>R^2 adjusted</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT</td>
<td>.785</td>
<td>.774</td>
</tr>
<tr>
<td>CU</td>
<td>.856</td>
<td>.832</td>
</tr>
</tbody>
</table>

4.3.4 Assessment for Structural Effect Size

By analyzing Cohen’s f^2, this study assessed the effect size of each path in the model. The f^2 values of 0.350 and above can show a large effect. And f^2 values between 0.150 and 0.350 can show a medium effect. After assessment, the findings are in Table 5.
Table 5: f-square

<table>
<thead>
<tr>
<th>Constructs</th>
<th>AT</th>
<th>CU</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT</td>
<td></td>
<td>0.835</td>
</tr>
<tr>
<td>CU</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG</td>
<td>0.312</td>
<td></td>
</tr>
<tr>
<td>HG</td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td>CO</td>
<td>0.498</td>
<td></td>
</tr>
<tr>
<td>EG</td>
<td>0.063</td>
<td></td>
</tr>
<tr>
<td>UG</td>
<td>0.202</td>
<td></td>
</tr>
</tbody>
</table>

5  CONCLUSIONS

5.1 Conclusion and Discussion

In this study, we investigated the gratification that drives attitudes and continuance use of mobile payment services in China, which is hardly mentioned by extant studies on consumer behavior studies in developing countries however, especially China. This study fills this gap and others indicated by this study by empirically applying U&G theory by collecting raw data on 200 individuals using mobile payment services. The results suggested that cognitive gratification, convenience gratification and usefulness gratification significantly impact users’ attitudes towards mobile payments. Furthermore, the findings show that Hedonic gratification and ease of use gratification did not impact user attitudes towards mobile payments. Similar to this study, the results of other studies have also revealed a positive effect of cognitive gratification on users’ attitude towards mobile payment \([6, 25]\). However, there are other study in another developing countries reporting opposite results \([1]\).

The results also suggest that Hedonic gratification can not significantly impact attitudes towards using mobile payments in China. As the results show, enjoyment users receive from mobile payment services may be not enough to change their attitudes. Other studies have also concluded that non-functional benefits have no significant effect in determining mobile payment user gratification \([6, 25]\). There are also studies that suggest that this hedonic non-functional benefit will trigger the user’s intrinsic desire to choose to use or not to use mobile payment services \([1]\). Different cultures may have created this difference, and in some cultures mobile payments exist only as a convenience. At the same time, users’ perceptions of mobile payment services themselves may have deepened this difference, and users could not find non-functional benefits in some functional services.

This study found that convenience gratification significantly affects Chinese users’ attitudes towards mobile payment services. This shows that users who use mobile payment services gain convenience from using mobile payment services, which also positively affects individuals’ attitudes towards mobile payment services. As demonstrated by the cognitive gratification already supported in this work, mature market environment, mature user habits, and mature product design, at the same time, encourage users to use mobile payment more. Among them, the long-term upgraded product design of Chinese service providers makes it
easy for users to feel the convenience gratification brought by mobile payment. This finding is consistent with studies by Ozturk et al. (2017) and Yonghee et al. (2016), who found that convenience gratification significantly affects user attitudes toward mobile payments. However, differences in culture, literacy levels, legislative and technological infrastructure may be important factors for the replication of the findings compared to studies in other developing countries [1, 22].

The hypothesis that the ease of use gratification of mobile payment services impacts user attitudes is rejected. This finding suggests that, in China, the ease of use of a mobile payment service does not affect whether users use it. Previous research argues that if users have access to information and knowledge about a mobile payment service as previously reported, they tend to find it easy to use [1]. Moreover, Mun et al. (2017) and Chuang et al. (2016) found that ease of use is a key factor affecting users’ attitudes towards using mobile payment services. However, in China, perhaps due to the high popularity of research on mobile payment services, users are not confused about how to use it, so a conclusion that is contrary to the previous research has emerged. This also raises the question of whether the key factors affecting users’ attitudes towards the mobile payment industry will be different at different stages of the development of the mobile payment industry.

This study showed that there is a positive correlation between usefulness gratification and mobile payment attitudes. Chinese users clearly recognize that mobile payment services can bring them better financial transactions, which has a positive impact on their mobile payment attitudes, leading them to continue using mobile payment services. Moreover, after years of technological development, mobile payment has created a series of trusted software stacks to provide users with full security. Such a trusted environment also helps to enhance users’ continuance use by gaining their trust.

To practice, the results can guide mobile payment service companies to improve their service capabilities by offering them users’ experience. Further, these findings will help to formulate relevant policies or rules for government, telecom and mobile payment companies to improve user gratification and loyalty. Based on the survey results, we recommend that mobile payment service providers enhance these desirable characteristics in the product design process in order to increase user’ continuance use.

5.2 Limitation and Future Research

However, this research was carried out on the basis of developing country China, its findings are difficult to generalize it to developed countries. At the same time, the developed mobile payment market in China also brings limitations to this study. Future research can further investigate the key factors impacting users’ continuance use of mobile payment services from the perspective of U&G theory by exploring the moderating role of culture and industry development stages.

REFERENCES


Temporal Fusion Transformers Model for Traffic Flow Prediction

Yuxuan Zhou
E-mail: ndzgbwdm@foxmail.com
The Hong Kong Polytechnic University, China

Abstract: Temporal Fusion Transformers (TFT) is a Transformer model for multi-step forecasting tasks. Because TFT models can integrate decoders to import various types of inputs, including static covariates, known future inputs, and other exogenous time series observed only in the past, which are well performed in the multi-step prediction of time series. To learn temporal relationships at different scales, TFT uses a cyclic layer for local processing and an interpretable self-attention layer for long-term dependence. TFT leverages specialized components to select relevant functions and inhibits unnecessary components through a series of gating layers to achieve high performance in a wide range of scenarios. When the model was proposed, it was considered to have good interpretability. As the research continues to increase, people put forward a lot of different opinions about this. This paper focuses on the explain ability of the TFT model and its attention mechanism.

Keywords: Transformer, Attention, Temporal Fusion Transformers model.

1 INTRODUCTION

Traffic flow is an important index to measure traffic conditions. Multi-step forward traffic flow prediction can help the traffic management agency to alleviate traffic congestion in advance and make the urban planning bureau more reasonable road planning [4]. It has a certain reference value for improving the prediction accuracy of multi-step leading speed [7]. However, accurate forecasting of traffic flows is particularly challenging due to the difficulty in capturing the spatio-temporal dependence of traffic data [5, 6].

Previous studies have mainly focused on various traditional statistical methods [8], such as vector autoregression-based (VAR) model and comprehensive autoregression-based moving average (ARIMA) model [9]. Recently, more and more neural network-based models have been applied to the field of traffic prediction [10]. Since the prediction of traffic flow is mainly based on the temporal correlation of traffic flow series, the circulating neural network (RNN) has been widely used to solve this problem. RNN can capture the time evolution of traffic speed, especially the long-time memory (LSTM) can learn the long-time dependence of speed series. Multiple RNN's were designed to predict [11]. However, the traditional RNN model has two problems: the huge demand for computing resources in the training process and the inability to consider long-term dependence [5]. To overcome the above shortcomings of RNN, researchers have begun to study
a new structure: the attention mechanism. Attention mechanism is widely used in the field of time series prediction because it can more effectively model dependencies sequentially [12].

In order to solve the above problems, this study adopts the transformer based time fusion transformer (TFT) structure to predict the traffic speed. Unlike the above methods, TFT is able to take into account a variety of input variables, and several new architectures have been introduced in TFT to improve predictive performance. (1) TFT uses gate control module and variable shielding network to fuse time information of velocity data of different scales together. The static information encoder is used to encode the number of detectors in the data acquisition field0 [2]. (2) TFT uses sequence-sequence layer to capture short-term time correlation in traffic speed time series, and uses self-attention mechanism to capture long-term time correlation in traffic speed time series. The research results of this paper are mainly reflected in two aspects: (1) When the prediction time is 30 min, the TFT model has a good prediction accuracy for multiple lanes. (2) In practice, the explainability of TFT models is difficult to determine.

2 DATES

![Figure 1: 31 sensor collection points](image)

The experimental data set is comes from the PeMS system of the transportation department of California, USA. In this experiment, I selected 31 sensor collection points in the San Jose Area and collected data for a total of 91 days from 2022/1/1 to 2022/4/1 for every 5mins. The data from January 1 to March 25 were selected as the training dataset to determine the model parameters and optimize the over parameters. The data from March 26 to April 1 were selected as the test dataset to conduct performance evaluation. Missing values are averaged using adjacent values. And abnormal value use the same time value of last week.
3 METHODS

3.1 Quantile output and loss functions

TFT supports quantile prediction. Let \( i \) represents unique entities in the traffic time series dataset. The definition of multi-step prediction problem can be simplified into the following formula:

\[
\hat{y}_i(q, t, \tau) = f_q(\tau, y_{i,t-k:t}, z_{i,t-k:t}, x_{i,t-k:t}, s_i)
\]  (1)

where, \( \hat{y}_i(q, t, \tau) \) under the point in time \( t \), the first to predict the future \( \tau \) on \( q \) points a numerical. \( f_q() \) is the prediction model. \( y_{i,t-k:t} \) : historical target variable. \( z_{i,t-k:t} \) : Time-varying variables that can be observed. \( x_{i,t-k:t} \) : Apriority-known Future Inputs. \( s_i \) : Static Covariates.

TFT generates point prediction intervals by simultaneously predicting different percentiles (e.g., 10, 50, and 90) for each time step. Quantile prediction is generated using the linear transform output of TFT’s decoder. Joint minimization of quantile loss was used to train TFT, and the outputs of all quantiles were added as follows:

\[
\mathcal{L}(\Omega, W) = \sum_{y \in \Omega} \sum_{q \in Q} \sum_{\tau=1}^{\tau_{\text{max}}} \frac{QL(y, \hat{y}_q(q, t - \tau, \tau), q)}{M_{t_{\text{max}}}}
\]  (2)

\[
QL(y, \hat{y}, q) = q(y - \hat{y})_+ + (1 - q)(\hat{y} - y)_+
\]  (3)

where, \( \Omega \) is the training data field containing samples, \( W \) represents the weight of TFT, \( Q \) is the set of output quantiles (\( Q = \{0.1, 0.5, 0.9\} \) used in the experiment). \( \mathcal{L}(\Omega, W) \) is the loss of quantiles \( q \) under the average prediction point of a single sequence. In this formula, due to \( (y - \hat{y})_+ \) and \( (\hat{y} - y)_+ \) will be one negative, one positive, so the formula can be converted into:

\[
QL(y, \hat{y}, q) = \max\left(q(y - \hat{y}), (1 - q)(\hat{y} - y)\right)
\]  (4)

In order to avoid the problem of inconsistent prediction dimensions under different prediction points, the author also did regularization processing, number 2 because only two quantiles of \( P50 \) and \( P90 \) are concerned here:

\[
q - \text{Risk} = \frac{2 \sum_{y \in \Omega} \sum_{\tau=1}^{\tau_{\text{max}}} QL(y_t, \hat{y}_t(q, t - \tau, \tau), q)}{\sum_{y_t \in \Omega} \sum_{\tau=1}^{\tau_{\text{max}}} |y_t|}
\]  (5)
4 MODEL ARCHITECTURE

Figure 2: The framework of Temporal fusion transformers model.

4.1 GRN (Gated Residual Network):
TFT consists of four main components, namely, gate mechanism, variable selection network, static covariate encoder, time processing and multilevel prediction interval prediction. The gating mechanism, which functions to skip all unused components of the architecture, provides adaptive depth and network complexity to accommodate different data sets and scenarios. The gated residual network (GRN) can make the nonlinear calculation between variables and targets of the model more flexible. GRN contains two types of input: primary input a and optional context c.

\[
GRN_\omega(a, c) = \text{LayerNorm}(a + \text{GLU}_\omega(\eta_1)) \\
\eta_1 = W_{1,\omega}\eta_2 + b_{1,\omega} \\
\eta_2 = \text{ELU}(W_{2,\omega}a + W_{3,\omega}c + b_{2,\omega})
\]

ELU is the Exponential Linear Unit activation function which is defined in follower equation:

\[
f(x) = \begin{cases} 
  x & \text{if } x > 0 \\
  \alpha(\exp(x) - 1) & \text{if } x < 0
\end{cases}
\]

\[
f'(x) = \begin{cases} 
  1 & \text{if } x > 0 \\
  f(x) + \alpha & \text{if } x < 0
\end{cases}
\]
\( \eta_1 \) and \( \eta_2 \) are intermediate layers; \( \eta_1, \eta_2 \in R^{d_{\text{model}}} \); \( \omega \) is an index to denote weight sharing. Letting \( \gamma \in R^{d_{\text{model}}} \) be the input the GLU is shown in following:

\[
\text{GLU}_\omega(\gamma) = \sigma(W_{\alpha,\omega}\gamma + b_{\alpha,\omega}) \odot (W_{\beta,\omega}\gamma + b_{\beta,\omega})
\]

(10)

Where \( \sigma(\cdot) \) is sigmoid activation function. \( W(\cdot) \in R^{d_{\text{model}} \times d_{\text{model}}} \), \( b(\cdot) \in R^{d_{\text{model}}} \) are the weights and biases. Via the GLU, GRN can control the structure of the model and neglect the unnecessary layers. Flexibility can be provided to inhibit any architecture that is not required for a given data set. Ensure the flow of effective information. The exact relationship between exogenous inputs and targets is often unknown in advance, making it difficult to foresee which variables are relevant. In addition, it can be difficult to determine how much nonlinear processing to do, and there may be situations where a simpler model will meet our needs -- for example, when the data set is small or noisy. In order to make the model flexibly apply nonlinear processing only when needed, we propose a controlled residual network. GLU can control the degree of nonlinear contribution.

4.2 VSN (Variable Selection Network)

The variable selects the network, and the corresponding input variable is selected at each time step. The variable selection network can screen out which variables are more important to the prediction problem and also remove all noise inputs in the TFT that may affect the prediction performance.

While multiple variables may be available, their correlation and specific contributions to the output are usually unknown. TFT is designed to provide instantiated variable selection by using a variable selection network applied to static and time-dependent covariates. In addition to providing insight into which variables are most important to the prediction problem, variable selection allows the TFT to remove any unnecessary noise inputs that might negatively affect performance. Most real-world time series data sets contain features with less predictive content, so variable selection can greatly aid model performance by leveraging learning capabilities only on the most significant features \(^3\).

Let \( \xi_t^{(j)} \in R^{d_{\text{model}}} \) be the input parameter after the transformation of the \( j \)-th variable at time \( t \) when \( \Xi_t = [\xi_t^{(1)^T}, \ldots, \xi_t^{(m_x)^T}]^T \) is the flattened vector of all past inputs at time \( t \). Putting \( \Xi_t \) and an external context vector \( c_t \) in GRN and then through a SoftMax layer can get the Variable selection weights \( v_{xt} \). So the equation is \( v_{xt} = \text{Softmax} \left( GRN_{v_{xt}}(\Xi_t, c_t) \right) \). The VSN section uses static, past and future inputs to select important features by the following equations:

\[
\xi_{st} = \sum_{i=1}^{m_x} \xi_t^i v_{xt}^{(i)} f_{s_t} (i)
\]

(11)

\[
\xi_t^{(i)} = GRN(\xi_{st}^{(i)})
\]

(12)
Different VSN is used for these three inputs which the parameters are not shared.

4.3 SCE (Static Covariate Encoders)

Different types of input variables should be treated differently. In this section, TFT is designed to generate four different context vectors, $c_e$, $c_c$, $c_h$. In fact, SCE is using GRN function. Those four context vectors are putting into different place in the TFD section(4). To be specific, $c_e$ is used in VSN, $c_c$, $c_h$ are devoted to initialize LSTM. And $c_e$ is used in SEL(Static Enrichment) layer in TFD section.

4.4 TFD (Temporal Fusion Decoder)

TFT refine multi-head attention in transformer-based architectures, TFT Refine multi-head attention in Transformer-based architectures, To enhance interpretability. It mainly has the following three major modules.

In time series data, important points are often identified based on the values around them, such as anomalies, points of change, or periodic patterns. TFT apply a sequence-to-sequence layer to reinforce those temporal relevance. Putting $\xi_{t-k_e}$ and $\xi_{t+1:t+t_{max}}$ in the LSTM encoder and decoder respectively. The input can be expressed as

$$\tilde{\phi}(t, n) = \text{LayerNorm}(\xi_{t+n} + \text{GLU}(\phi(t, n)))$$

$$\phi(t, n) \in \{\phi(t, -k), \ldots, \phi(t, t_{max})\}$$

n is the position index. And $c_e$, $c_h$ from 4.3 are devoted to initialize the cell state and hidden state respectively for the first LSTM in the layer.

4.4.1 SEL (Static Enrichment Layer)

Since Static information usually has a significant impact on the accuracy of time series prediction, Static Enrichment Layer enhances timing feature by introducing static covariable, namely simply using GRN and input $c_e$ given by the static covariable encoder.

$$\theta(t, n) = \text{GRN}_\phi(\tilde{\phi}(t, n), c_e)$$

4.4.2 TSL (Temporal Self-Attention Layer)

The self-Attention module can learn long-term dependencies on time series data and provide for model interpretability. In TSL, it is mainly the interpretable polycephalic self-concern layer, plus GLU. The interpretable multi-head attention used in every moment ($N = t_{max} + k + 1$) where $\Theta(t) = [\theta(t, -k), \ldots, \theta(t, t)]^T$

$$B(t) = \text{InterpretableMultiHead}(\Theta(t), \Theta(t))$$

$$\theta(t, \Theta(t))$$
\( \delta(t,n) = \text{LayNorm}(\theta(t,n)) + \text{GLU}(\beta(t,n)) \)  \hspace{1cm} (17)

To be more specific, the Interpretable multi-head attention has three major parameters: query, key, and value and then uses the scale dot-product as follows:

\[
\text{Attention}(Q, K, V) = A(Q, K)V \hspace{1cm} (18)
\]

\[
Q \in R^{N \times d_{\text{attn}}}, K \in R^{N \times d_{\text{attn}}}, V \in R^{N \times d_{V}} \hspace{1cm} (19)
\]

Where \( A(\cdot) \) is a normalization function, and \( N \) is the time steps.

\[
(Q, K)V = \text{Softmax}(QR^T \sqrt{d_{\text{attn}}}) \hspace{1cm} (20)
\]

For the multi-head aspect, this mechanism employs different heads for different representation subspaces:

\[
\text{MultiHead}(Q, K, V) = [H_1, ..., H_m]W_H \hspace{1cm} (21)
\]

\[
H_h = \text{Attention}(QW_Q^{(h)}, KW_K^{(h)}, VW_V^{(h)}) \hspace{1cm} (22)
\]

Where \( W_Q^{(h)} \in R^{d_{\text{model}} \times d_{\text{attn}}}, W_K^{(h)} \in R^{d_{\text{model}} \times d_{\text{attn}}}, W_V^{(h)} \in R^{d_{\text{model}} \times d_{V}}, W_H \in R^{(m_H \times d_V) \times d_{\text{model}}}, W_H \) is the Linear combinations of all the heads \( H_h \).

Attention weights separately can’t indicate the importance of a particular feature, so TFT modifies multi-head attention to share values in each head and employ additive aggregation of all heads. In other words, for \( V \) is the multi-head shared parameter, for \( Q \) and \( K \) are the multi-head independent parameters, then calculate the multi-head attention score weighted \( V \), and sum the average output.

\[
\text{InterpretableMultiHead}(Q, K, V) = \bar{H}W_H \hspace{1cm} (23)
\]

\[
\bar{H} = A(Q, K)VW_V = \left\{ \frac{1}{m_H} \sum_{h=1}^{m_H} A\left(QW_Q^{(h)}, KW_K^{(h)}\right) \right\} V W_V
\]

\[
= \frac{1}{m_H} \sum_{h=1}^{m_H} \text{Attention}(QW_Q^{(h)}, KW_K^{(h)}, VW_V) \hspace{1cm} (24)
\]
4.4.3PFL (Position-wise Feed-forward Layer)

Apply additional nonlinear processing to the output of the self-focused layer.

\[
\psi(t, n) = GRU_{\phi}(\delta(t, n))
\]

InterpretableMultiHead\((Q, K, V) = \bar{H} \, W_H\)

\[
\bar{\psi}(t, n) = LayerNorm(\tilde{\psi}(t, n) + GLU_{\phi}(\psi(t, n)) \right)
\]

5 RESULTS

Table 1: Hyperparameter search ranges.

<table>
<thead>
<tr>
<th>Hyperparameter</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>State size</td>
<td>10, 20, 40, 80, 160, 240, 320</td>
</tr>
<tr>
<td>Dropout rate</td>
<td>0.1, 0.2, 0.3, 0.4, 0.5, 0.7, 0.9</td>
</tr>
<tr>
<td>Minibatch size</td>
<td>20, 30, 40, 50, 64, 128</td>
</tr>
<tr>
<td>Learning rate</td>
<td>0.001, 0.002, 0.004, 0.008, 0.01</td>
</tr>
<tr>
<td>Heads</td>
<td>1, 2, 3, 4</td>
</tr>
</tbody>
</table>

Bold values indicate the optimal hyperparameters.

Figure 3: Prediction results of the TFT on lane 1
Figure 4: Attention score of TFT on lane 1

Table 2: Prediction results of different models.

<table>
<thead>
<tr>
<th>Model 1</th>
<th>Model 1-max</th>
<th>Model 1-min</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE</td>
<td>122.4416</td>
<td>1112.1327</td>
</tr>
<tr>
<td>RMSE</td>
<td>190.3704</td>
<td>1189.4219</td>
</tr>
<tr>
<td>MAPE</td>
<td>0.1791</td>
<td>1.9337</td>
</tr>
<tr>
<td>R² score</td>
<td>0.8751</td>
<td>-3.8746</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model 2</th>
<th>Model 2-max</th>
<th>Model 2-min</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE</td>
<td>175.5456</td>
<td>498.4471</td>
</tr>
<tr>
<td>RMSE</td>
<td>276.9062</td>
<td>643.5049</td>
</tr>
<tr>
<td>MAPE</td>
<td>0.1816</td>
<td>0.3482</td>
</tr>
<tr>
<td>R² score</td>
<td>0.8723</td>
<td>0.3103</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model 3</th>
<th>Model 3-max</th>
<th>Model 3-min</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE</td>
<td>148.3931</td>
<td>170.6433</td>
</tr>
<tr>
<td>RMSE</td>
<td>199.7648</td>
<td>218.4773</td>
</tr>
<tr>
<td>MAPE</td>
<td>0.1483</td>
<td>0.1768</td>
</tr>
<tr>
<td>R² score</td>
<td>0.9491</td>
<td>0.9392</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model 4</th>
<th>Model 4-max</th>
<th>Model 4-min</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE</td>
<td>132.7735</td>
<td>156.7910</td>
</tr>
<tr>
<td>RMSE</td>
<td>180.2207</td>
<td>197.7319</td>
</tr>
<tr>
<td>MAPE</td>
<td>0.1379</td>
<td>0.1738</td>
</tr>
<tr>
<td>R² score</td>
<td>0.9537</td>
<td>0.9442</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model 5</th>
<th>Model 5-max</th>
<th>Model 5-min</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE</td>
<td>126.5685</td>
<td>151.6295</td>
</tr>
<tr>
<td>RMSE</td>
<td>277.9803</td>
<td>305.3027</td>
</tr>
<tr>
<td>MAPE</td>
<td>0.0956</td>
<td>0.1200</td>
</tr>
<tr>
<td>R² score</td>
<td>0.9252</td>
<td>0.9098</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model 6</th>
<th>Model 6-max</th>
<th>Model 6-min</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE</td>
<td>122.4416</td>
<td>100.8592</td>
</tr>
<tr>
<td>RMSE</td>
<td>190.3704</td>
<td>157.5504</td>
</tr>
<tr>
<td>MAPE</td>
<td>0.1791</td>
<td>0.1273</td>
</tr>
<tr>
<td>R² score</td>
<td>0.8751</td>
<td>0.9145</td>
</tr>
</tbody>
</table>
6 CONCLUSIONS

In order to reduce the uncertainty of prediction results, for each lane, TFT was run 10 times and the model with the smallest loss was selected. Then, the first and last lanes with the highest importance of encoder variables were excluded respectively for comparison experiment. Each run of the model predicts a different outcome. That is because deep learning is a random machine learning algorithm. On the one hand, when training the neural network, the weight of the neural network is initialized and random. On the other hand, this study sets the exit rate in the algorithm to avoid overfitting. Setting the exit rate means that the algorithm will randomly discard some neurons and network nodes during training. It can be seen from the test results that in multiple modelling (more than 300 times), the performance of the model is completely different, and its loss and attention score are also completely different. Further analysis shows that after removing the lane data with the most and least weight, the loss change of the model cannot be analysed and it is difficult to find out the rule. The predictive power of its models also fails to find patterns. Here, I think the main reason is the randomness of deep learning. In multiple learning, the parameters of the model obtained each time are completely different, which is difficult to reproduce. Therefore, it is difficult to reproduce the attention score of each model. To sum up, the explanatory power of its model cannot be demonstrated stably for more than 300 hours. In theory, the attention mechanism should be able to express certain explanatory ability, but in practice, it is impossible to know whether the model calculated next time will be better and more suitable to the data. The calculation difficulty is too complicated, and the calculation power needs are too large.

In many experiments, the TFT model has high predictive performance and can better simulate the time correlation through the self-attention mechanism, which provides experience for the establishment of a good long-term forecast model. And the explanatory prediction process can be given. However, it is difficult to determine whether this model is an optimal model or an interpretable model. On the other hand, TFT requires a lot of data to train the model to achieve good predictive performance and requires a lot of time and computing power. Since graph neural network can capture spatial correlation, TFT and graph neural network can be combined to predict the traffic state of road network. Another interesting area of research is the integration of traffic data, geographic information, weather information and other information into speed forecasting.

REFERENCES

Design and Development of University Smart Campus Platform Based on Big Data

Daigen Huang
543948894@qq.com
Sichuan University of Culture and Arts, Mianyang, Sichuan, China

Abstract: Under the background of the in-depth development of digital society, the informatization process of colleges and universities will gradually change to intelligence and intelligence, and the new generation of information technology represented by big data technology has become the core force of building a "smart campus". In this regard, this paper takes the overall architecture of "Smart Campus" as the research object, and comprehensively integrates the practical characteristics of big data technology, network information technology and computer application technology. On the one hand, Hadoop cluster is used to build a data analysis and processing server. On the other hand, we will design a Web-based intelligent campus application service platform based on J2EE technical specifications. The whole platform adopts B/S architecture design, and the Web Server completes the control of various business logics and the configuration of corresponding API data interfaces, which facilitates users to query and call many data resources such as teaching, scientific research, student behavior, campus management, etc. through simple interactive operation of front-end pages. It also relies on data mining algorithm models such as K-means and Person to complete the analysis and processing of massive data, which optimizes the teaching management mode and improves the level of educational services in colleges and universities, thus making benefits for the construction of smart campus in colleges and universities.

Keywords: Big Data, Smart Campus, Hadoop, Data Analysis, Computer Application.

1 INTRODUCTION

With the release of China's "14th Five-Year Plan" and the outline of the 2035 long-term goal, building a high-quality education system in an all-round way has become the starting point and focus of further deepening the reform and innovative development of China's higher education, and it is also the only way to meet the needs of education modernization and achieve a strong education country. Under the high-quality education system, the concept, culture and ecology of education are constantly innovating and changing. Especially under the influence of the new generation of information technologies such as network communication, big data and artificial intelligence, the education mode and talent training objectives have undergone unprecedented changes, and the development of educational informatization has entered a critical period of transformation and upgrading. [6]

At present, the deployment of information infrastructure in colleges and universities is becoming more and more perfect, and various education and teaching systems and campus
management service systems are everywhere. However, the original intention of most designs and constructions is based on the concept of "on demand, independence and one by one", so that the data application caliber among various systems is not uniform, and there is a lack of data correlation and interaction, thus forming an isolated island structure of data information, hardware equipment and application services, which not only increases the application difficulty of the system, but also makes the management and operation and maintenance of the system complicated and chaotic. [10] Traditional and single information means can't meet the development requirements of campus comprehensive management and knowledge resource service in the era of big data and artificial intelligence. However, the proposal of smart campus can fully perceive the campus environment with the support of big data technology, and highly unify and correspond the physical space and data space, fully contain all kinds of characters, events and tasks and data information on campus, and manage and monitor all kinds of information and data on campus in various directions by using data modeling and data mining technologies, so as to realize the reconstruction of campus development ecology. In view of this, this paper holds that colleges and universities should adhere to the strategy of educational modernization and high-quality development driven by educational informatization, accelerate the innovation and integration of the new generation of information technology and the practice of educational and teaching activities, give full play to the application advantages of big data technology, network information technology and computer application technology, and build a smart campus application service platform in colleges and universities. Based on the interactive application of data processing center and visualization, the platform of "Smart Campus" comprehensively covers three major sections: teaching, scientific research and daily management, completes the functional integration of each subsystem, enhances the comprehensive utilization of data resources, improves the service level of campus, and makes a useful attempt to construct a paradigm of smart education in colleges and universities.

2 INTRODUCTION OF KEY TECHNOLOGIES

2.1 Big Data Technology

Big Data generally refers to the data collection whose data volume has exceeded the normal range, and users can't use ordinary software tools and methods to process it. Big data comes from the accumulation of network communication technology, which is the inevitable result of the application of various digital technologies. Since big data was put forward, it has formed a huge and dynamic concept after years of development. On the one hand, it represents a huge amount of data information resources, on the other hand, it represents the general name of a series of data value technologies. [9] Big data technology is the foundation of data value, and it is also the key to the transition of big data from single data level to application level. Big data technology is a collection of technologies covering the whole life cycle of big data, covering a series of links such as data collection, transmission, cleaning, storage, analysis, processing, presentation and application. It provides a complete processing paradigm for the acquisition and reflection of the value of big data, and also expands the application scenarios of big data. With the wide application of big data, the development speed of big data technology is constantly accelerating, and many technical means are gradually gathering and forming a systematic and ecological big data technology stack.
2.2 Hadoop Technical Framework

Hadoop is an open source technology framework realized by Java language, and it is also the most widely used big data core technology at present. Hadoop is a distributed system architecture, and its interior is mainly composed of HDFS (Hadoop Distributed File System) and MapReduce, which can be used as the core of the underlying storage and analysis to provide users with reliable, scalable and distributed computing big data services. In addition, the high scalability of Hadoop framework can support the deployment and installation of various functional components, and they are compatible with each other, forming an independent application system and gradually evolving into Hadoop ecosystem, as shown in Figure 1. Hadoop supports cluster deployment, which is convenient for users to quickly complete the design and development of large-scale data analysis applications, and realize the parallel high-speed operation and complex call of big data.

![Figure 1: Hadoop ecosystem (original)](image)

2.2.1 HBase

HBase is a distributed database with column storage, but HBase itself is not directly involved in file storage, and its actual functions are still realized by HDFS under Hadoop framework. The design core of HBase is to realize random and real-time read/write access of HDFS system. As the representative of non-relational database, its main functional modules include HMaster, Region Server and ZooKeeper. Among them, HBase supports users to complete various operations in HBase database through Java API under HBase Client, and HMaster is responsible for the distribution and management of Region Server, realizing the overall load balance of the system and the control of various permissions.

2.2.2 Spark

Spark is a more efficient distributed parallel computing framework. Compared with MapReduce, the whole calculation process of Spark framework is based on memory, which greatly reduces the reading and writing time of disk, and enables RDD (Resilient Distributed Datasets) as a new data structure to realize a more brief and concurrent calculation process. Spark framework contains four parts: tool layer, calculation layer, storage layer and resource scheduling layer, in which the calculation layer (Spark Core) is the running foundation of the whole framework, which is mainly responsible for the scheduling of distributed computing tasks and the basic control of input and output interfaces. In addition, the Spark framework will be extended based on Spark Core, integrating four core components of Spark SQL, Spark
Streaming, MLlib and GraphX to enrich the application scenarios of Spark framework in different fields.

2.3 JavaWeb

JavaWeb is the sum total of technologies that use Java language to complete the development of dynamic Web applications, which can help developers quickly solve various problems of Web client or Web server. JavaWeb is subordinate to J2EE technical specification, which can divide Web application into four parts: presentation layer, control layer, business logic layer and persistence layer, and support the development and deployment of each part by JSP page, Servlet, JavaBean or EJB and JDBC. In addition, with the continuous expansion of the scale of Web applications, the related functions and requirements are becoming more and more complex, and JavaWeb has entered the framework stage. Many frameworks represented by Struts, Spring, Hibernate and Mybatis have been integrated and changed to form SSH or SSM, which can greatly speed up the development process, reduce time and cost, and realize the agile development of web applications.

2.4 Data Mining

Data Mining (DM) is a process of extracting hidden information with certain potential value from a large number of disordered data information with the help of computer application technology. The essence of data mining is to build a data analysis and processing model, taking data information as the research object, and providing help for subsequent decision-making through analysis and prediction. The process of data mining can be simply divided into four steps: problem definition, data preparation, data mining and result analysis. Among them, the choice of data mining algorithm is the core of the whole data mining work, and it is also the key to build the corresponding data mining model. Common data mining methods include classification analysis, prediction analysis, cluster analysis, valuation analysis and correlation analysis, as shown in Table 1. As for the final result analysis, it is more inclined to the correlation between data, the trend and trend of data and the characteristic expression of specific objects.

Table 1: Type of common data mining algorithms (original)

<table>
<thead>
<tr>
<th>Classify</th>
<th>Data mining method</th>
<th>Data mining model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guided data mining</td>
<td>Classification analysis</td>
<td>Decision Tree, Random Forest, Neural Network</td>
</tr>
<tr>
<td></td>
<td>Predictive analysis</td>
<td>Regression tree, Rough Set method</td>
</tr>
<tr>
<td></td>
<td>Valuation analysis</td>
<td>SVM, Bayesian method</td>
</tr>
<tr>
<td>Unsupervised data mining</td>
<td>Correlation analysis</td>
<td>Pearson correlation coefficient, Aprior algorithm</td>
</tr>
</tbody>
</table>

2.5 Development Process

According to the application requirements of the above related application technologies, complete the configuration and deployment of the development environment of the application service platform of smart campus in colleges and universities. The development content of the system is divided into two parts. One is to build Hadoop cluster, complete the collection and storage management of all kinds of campus data information, and build a data analysis and
mining engine with the help of Spark framework. Secondly, under the Java development environment, the SSH framework is used to complete the development of Web Server, and the development and deployment of various lightweight service bus interfaces are completed, thus forming a standard Web application.

First of all, Hadoop cluster architecture needs the support of hardware and software. The underlying operating system is Linux, CentOS 6.7(x86_64) is the version, and jdk-8u291-linux-x64 is the JDK version. According to the application requirements of the system, Hadoop will be deployed in a completely distributed cluster. There are 7 nodes in the cluster, which are named Master1, Master2, Slave1, Slave2, Slave3, Slave4 and Slave5 respectively. Hadoop is version 2.7.7, which is installed in each node, and components such as Yarn, HDFS, Zookeeper, HBase, Sqoop and Kafka are also deployed in each node. Among them, Sqoop component can import all kinds of structured data into HDFS system under Hadoop architecture to realize distributed storage and form raw data. Kafka can import the operation logs and unstructured data of the system into HBase for storage.

Secondly, for the data mining analysis function, the system will use Spark framework to read all kinds of data in HDFS. After data preprocessing, it will complete the data call processing by constructing the corresponding data mining model. As shown in Figure 2, the K-means clustering algorithm code is implemented for Spark.

```scala
object KMeans {
  def main(args: Array[String]) = {
    val conf = new SparkConf().setAppName("K-Means").setMaster("spark://master:5077")
    .setJarByClass(SimpleGraphXJar.getClass)()
    val sc = new SparkContext(conf)
    val data = sc.textFile("hdfs://master:9000/kmeans_data.txt", 1)
    val parsedData = data.map(_.split(" ")\(map\_todouble\))
    val numClusters = 2
    val numIterations = 20
    val model = KMeans.train(parsedData, numClusters, numIterations)
    println("Cluster centres")
    for (c <- model.clusterCenters) {
      println("\" + c.toString + \"")
    }
  }
}
```

**Figure 2: K-means algorithm model building code (original)**

Finally, for the development of Web application server, the basic development language is Java, MyEclipse 2018 is the integrated environment, Tomcat 8.0 is the Web server, and MySQL 5.7 is the database server. And the project object model (Maven) is used to manage the project structure. Maven chooses Apache-Maven-3.2.1 version. In the process of building the overall development environment, the installation of JDK and the configuration of environment variables are completed first to build the foundation of Java application development. Secondly, the installation of MyEclipse and the installation of Tomcat, the Web server, and the configuration of Tomcat is completed in the Preference option under MyEclipse. Then, based on SSH architecture, complete the integration and encapsulation of the whole system. Through the introduction of the above key technical theories, the overall environment of the system development, the configuration of related software and tools are
determined, and the technical feasibility of the overall project of the application service platform of smart campus in colleges and universities is also clarified.

3 FUNCTION REALIZATION

3.1 Data Management

The platform has a unified initialization login interface, which makes the previous decentralized education and teaching system and campus management service system business platform integrated. The platform uses big data technology to establish a perfect data sharing center, collect the historical data of various existing systems on campus, complete the lossless transplantation and optimized integration of data resources, and realize the distributed storage of data information with Hadoop cluster to realize the comprehensive management of campus data. [4] Under the data management function module, the system will give different data usage rights according to different user roles. For example, student users can independently check the semester and academic year plans, keep abreast of the learning progress, and know their own credits and grades. Teacher users can quickly handle all kinds of administrative, party-mass and other daily work, realize the joint search and use of data and information, and effectively improve work efficiency.

3.2 Teaching Service

According to the actual application needs of students and teachers, the platform will set up two sub-functions of education and teaching management and scientific research achievement management under the teaching service module. Under the function of education and teaching management, teacher users can upload teaching materials, publish relevant teaching plans and curriculum tasks, arrange online homework and tests, and make statistical analysis of students' learning situation. The functions of student users include online learning, data downloading, online homework and testing.

As for the management function of scientific research achievements, teacher users have the right to complete scientific research project application, scientific research project audit, personnel information query, scientific research projects and achievements query, etc. in the platform, which can reduce a lot of complicated and trivial information collection, improve efficiency and ensure quality.

3.3 Campus Management

As the core application of "Smart Campus" platform, it will take students' daily behavior on campus as the research object, conduct multidimensional data analysis and mining, and provide scientific decision-making basis for the management of various affairs on campus. Taking the student behavior analysis model as an example, the data features are selected from five dimensions, including classroom learning, online learning, other ways of learning, campus life and daily entertainment, and the corresponding data sources come from the internal database of the platform, campus video surveillance system and campus "all-in-one card" system. As shown in Table 2, the information table of students' behavior characteristics, the platform will synthesize all kinds of characteristics to construct students' behavior patterns, that is, student user portraits.
Table 2: Information table of students' behavior characteristics (original)

<table>
<thead>
<tr>
<th>No.</th>
<th>Data dimension</th>
<th>Data features</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Classroom learning</td>
<td>Course characteristics</td>
<td>Course name, course teacher</td>
</tr>
<tr>
<td>2</td>
<td>Performance characteristics</td>
<td>Normal grades, comprehensive grades</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Online learning</td>
<td>Learning duration characteristics</td>
<td>System usage time</td>
</tr>
<tr>
<td>4</td>
<td>Online learning</td>
<td>Job completion characteristics</td>
<td>Job completed, job not completed</td>
</tr>
<tr>
<td>5</td>
<td>Other ways to learn</td>
<td>Book borrowing characteristics</td>
<td>Borrowing time, book name</td>
</tr>
<tr>
<td>6</td>
<td>Campus life</td>
<td>Characteristics of dining in canteen</td>
<td>Location, amount</td>
</tr>
<tr>
<td>7</td>
<td>Campus life</td>
<td>Shopping characteristics</td>
<td>Time, amount</td>
</tr>
<tr>
<td>8</td>
<td>Daily entertainment</td>
<td>Participation characteristics</td>
<td>Activity time, activity name</td>
</tr>
</tbody>
</table>

Through the user portrait, data mining analysis can be carried out for the corresponding problems. For example, in the analysis of association rules between students' behaviors and students' achievements by Apriori algorithm, the calculation formulas of association rules are shown in formulas 1 and 2. $S$ stands for support, that is, the proportion of student behavior (X) and student achievement (Y) in all data sets; And $C$ stands for confidence, that is, the ratio of student behavior (x), student achievement (y) and student behavior (x) support. (Gao, 2022)

The results of the final calculation are shown in Table 3, and the platform can also display the corresponding data results in the form of charts, as shown in Figure 3, which is a scatter plot of the correlation between achievement and shopping characteristics.

\[
S(X \Rightarrow Y) = \frac{\text{count}(X \cup Y)}{|D|} \quad (1)
\]

\[
C(X \Rightarrow Y) = \frac{S(X \cup Y)}{S(X)} \quad (2)
\]

Table 3: calculation formula for some indexes (original)

<table>
<thead>
<tr>
<th>No.</th>
<th>Consequent</th>
<th>Antecedent</th>
<th>Support (S)</th>
<th>Confidence (C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Good grades</td>
<td>Learning duration characteristics</td>
<td>19.77</td>
<td>86.55</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Book borrowing characteristics</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Participation characteristics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Poor grades</td>
<td>Characteristics of dining in canteen</td>
<td>17.69</td>
<td>66.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Participation characteristics</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Course characteristics</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4 CONCLUSIONS

With the aim of building a "smart campus" in the new era, this paper builds an application service platform for smart campus in colleges and universities based on the functional characteristics of big data technology, network information technology and computer application technology. The function of the platform combines the current campus education and teaching systems and campus management service systems, completes the sharing and interaction of all kinds of data, and provides a brand-new data analysis system. It is convenient for users to realize the application and management of teaching, scientific research and daily management with concise, convenient and efficient operation, which improves the comprehensive utilization rate of data resources, improves the service level of campus, and makes a beneficial attempt to construct the wisdom education paradigm in colleges and universities.

REFERENCES

Research on Precision Marketing Strategy Based on Clustering Algorithm

Zhengmin Fu
767483891@qq.com
University of Chinese Academy of Social Sciences, Beijing, China

Abstract: With the development of personalized services, the use of big data technology to guide precision marketing has become a trend in the future development of e-commerce. However, data mining algorithms such as clustering algorithms to parse precision marketing patterns have not been widely studied. In this paper, a python crawler and a relevant public dataset were used to collect 25,000 data from a shoe store in Taobao, and the data composition came from two aspects: user attribute data and transaction data. The data were pre-processed using SPSS software, and the RFM model was established and standardized, then the relevant value weight coefficients were derived using Matlab software, and the data were analyzed by clustering algorithm using SPSS, and the total customer value was used to verify the clustering results. Combining the results of text analysis and clustering algorithm analysis, we finally propose a precise marketing strategy.

Keywords: Clustering Algorithm, RFM Model, Total Customer Value, Precision Marketing.

1 INTRODUCTION

From the previous literature, it can be seen that the concept of precision marketing and the concept of data mining appeared relatively early, and most scholars have a wide range of research on the impact of data mining on precision marketing, but there is still relatively little research on the application of data mining technology to e-commerce precision marketing, and most scholars have only studied what kind of impact precision marketing in the context of big data can bring to the development of e-commerce, but for how to Most scholars have only studied the impact of precision marketing on the development of e-commerce in the context of big data. [1]. The rapid popularity of the Internet has brought new opportunities for the development of e-commerce industry, and the rapid rise of big data in recent years has also provided new impetus for the development of e-commerce industry, and it can be seen from the development in recent years that big data technology has played a huge role in many fields. How to take advantage of the big data background in the field of e-commerce, through data mining technology to help the e-commerce industry to better identify target customers and achieve accurate marketing is the main topic of this paper focus on research. With the development of society, the traditional seller's market has been gradually replaced by the buyer's market, and consumers are more inclined to personalization when purchasing products. This demand for personalization now extends to the marketing field, where consumers want to get the marketing ads they want to see. The e-commerce industry uses data mining techniques to analyze user
behavior data, and this information will help e-commerce platforms to implement more accurate marketing and advertising strategies. [2] This paper provides practical reference value for e-commerce companies to use big data technology to analyze the precise marketing model through case study analysis. [3]

2 MODEL DESIGN AND DATA ANALYSIS

2.1 Model Framework Design

(1) Collect user-related transaction data of e-commerce industry, pre-process the data, remove the abnormal data in the data set, and then standardize the pre-processed data.

(2) The values of R, F and M are calculated from the data and constitute the value matrix of users.

(3) Apply the clustering algorithm to cluster the matrix and derive the corresponding user categories.

(4) Apply principal component analysis to calculate the importance of each indicator in the matrix as the weight value of the indicator.

(5) Use the weight values calculated in step 4 to calculate the value of each category and each user.

(6) The value of users is used to verify the effect of clustering.

(7) For each category, propose a specific marketing strategy based on the characteristics of the user to achieve the purpose of precision marketing.

![Marketing Model Design](image)

**Figure 1:** Marketing Model Design
2.2 Data Processing

The first point is that there are a large number of missing values in some indicators in the user attribute data and sales data; the second point is that there are refunded users in the sales data, which is not meaningful for user value calculation and will have a great impact on the model results if this part of data is kept. When further analyzing the user attribute data, we can see that the missing values of consumption level and city level indicators account for about 53.9% and 32.6% respectively. Since the city level indicator is a categorical indicator, this paper uses the plural filling method for filling. [4]

After cleaning and sorting the data set, the indicators of RFM model need to be constructed. According to the introduction of Chapter 2, it is known that: R indicator is the time when the customer purchased the goods in the previous time; F indicator is the total number of times the customer's consumption occurred from January 2021 to June 2021; M indicator is the average consumption amount of the customer from January 2021 to June 2021; using the user consumption dataset to calculate according to the definition of the three indicators. [5]

2.3 Calculation of RFM model indicators

In this paper, principal component analysis is used to calculate the weight value of each indicator. The principal component analysis method is a method to simplify the data set. It mainly uses orthogonal transformation to linearly transform all possible relevant variables to obtain a series of linearly uncorrelated variables. These uncorrelated variables derived from the original variables by linear transformation from are called principal components. [6] The specific implementation steps of principal component analysis are as follows. [7]

(1) Centering the original data. Centering is to have each data point subtracted from the mean of the category to which it belongs.

(2) Derive the covariance matrix of the features. If the original data has n features, the covariance matrix is a matrix of order n.

(3) The covariance matrix is decomposed by the eigenvalues. The eigenvalues and eigenvectors of the above covariance matrix are calculated. In this paper, if there are three indicators R, F and M, then 3 eigenvalues will be obtained, and each eigenvalue corresponds to one indicator. Expressed in the formula as follows.

\[ \gamma \alpha = \Gamma \alpha \]

where \( \alpha \) denotes the matrix composed of eigenvectors and \( \gamma \) denotes the column vector composed of eigenvalues.

(4) The eigenvalues corresponding to each indicator are used as the weights of each indicator. We calculate the value of each user, which is calculated by the formula

\[ V = \alpha_1 \times R + \alpha_2 \times F + \alpha_3 \times M \]
where \((\alpha_1, \alpha_2, \alpha_3)\) is the eigenvalue derived by principal component analysis and V is the user's value matrix. After clustering, the value matrix of each category of users is then calculated according to the above formula, and the value matrix before clustering is used to verify the value matrix after clustering.

The principal component analysis method introduced above is used to carry out the weight value calculation of each indicator. Usually, the hierarchical analysis method is used to calculate the weight value of each indicator of the RFM model, which requires scoring the R/F/M value given to each consumer and then constructing a judgment matrix based on the score of each indicator. However, the hierarchical analysis method has less quantitative data and more qualitative components, which is not easily convincing, while when there are too many indicators, the data statistics are large and the weights are difficult to determine. The principal component analysis method is to take the indicators with linear correlation and recombine them, so as to obtain a new set of linearly unrelated composite indicators to replace the original ones. [8]

### 2.4 Data Clustering

The main steps of the experiment are.

1. install the Python tool;
2. import the processed data into the dataset and set the title for the dataset;
3. import the K-Means analysis tool into the Python environment;
4. normalize the imported data;
5. build the analysis model and set the clustering K value;
6. load the feature vector to be analyzed and perform the algorithm operation;
7. obtain the classification results and merge them with the dataset;
8. output Final results. The main procedure is as follows.

```python
# import the K-Means analysis tool into the python environment and import the dataset using pandas
from sklearn.cluster import KMeans
import pandas as pd

#Read the data used for clustering and create the teaching data table
loan_data=pd.DataFrame(pd.read_csv('SJRY_data.csv',header=0))
loan_data.columns
Index(['XM','ZW','GT','WT','age','DT','zc','ZCJB','XL','ZY','JZJ'],dtype='object')

#Process the training set and normalize the analysis data features
loan_data_zs=1.0*(loan_data-loan_data.mean())/loan_data.std()
```
# build the analytical model, load the feature vector "job title, time in the workforce and age" that needs to be clustered, the number of clusters is 3
loan=np.array(loan_data[['ZW','WT','age']])
clf=KMeans(n_clusters=3)
#Substitute the data into the clustering model
clf=clf.fit(loop)
#Perform simulation training to get predicted values
Cluster=clf_KMeans.fit_predict(X)
print(cluster)
#Add clustering result labels to the original data table
loan_data['label']=clf.labels_
#View the clustering results
print(loan_data)

2.5 Cluster feature analysis
The clustered dataset was analyzed and the number of users in each clustering category is shown in Table 1.

<table>
<thead>
<tr>
<th>Clustering</th>
<th>Number of cases in each cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5199.000</td>
</tr>
<tr>
<td>2</td>
<td>8474.000</td>
</tr>
<tr>
<td>3</td>
<td>4307.000</td>
</tr>
<tr>
<td>4</td>
<td>7020.000</td>
</tr>
<tr>
<td>Effective</td>
<td>25000.000</td>
</tr>
<tr>
<td>Missing</td>
<td>.000</td>
</tr>
</tbody>
</table>

The mean, maximum and minimum values of each indicator were calculated for each category, as shown in Table 2.

<table>
<thead>
<tr>
<th>Indicators</th>
<th>Category</th>
<th>The first category</th>
<th>The second category</th>
<th>The third category</th>
<th>The fourth category</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>Average value</td>
<td>38.09</td>
<td>143.91</td>
<td>37.13</td>
<td>142.57</td>
</tr>
<tr>
<td></td>
<td>Maximum value</td>
<td>171</td>
<td>181</td>
<td>174</td>
<td>181</td>
</tr>
<tr>
<td></td>
<td>Minimum value</td>
<td>1</td>
<td>8</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>F</td>
<td>Average value</td>
<td>1.34</td>
<td>1.24</td>
<td>1.33</td>
<td>1.15</td>
</tr>
</tbody>
</table>
From the above table, it can be seen that the R indicator reflects the time of the customer's last consumption from now, the customers of the first and third category have a shorter time from the last consumption, which means that the higher the trust of these two categories of customers to the enterprise, proving that these two categories of customers are quality customers, but the values of the second and fourth categories are larger, and from the maximum value, we can conclude that some customers have not made repeat purchases in a long time cycle. the F indicator is M indicator refers to the average amount of money spent by customers over a period of time, and the higher value of the average consumption of customers in the first and second categories indicates that they contribute more to the profit of the company and may be loyal customers of the store.

To further measure each indicator more accurately for different categories of users so that the categories of users can be identified, the average value of each indicator is introduced and the results are shown in Table 1.

| M | Maximum value | 7 | 7 | 7 | 7 |
| Minimum value | 1 | 1 | 1 | 7 |
| Average value | 187.48 | 181.89 | 63.17 | 61.96 |
| Maximum value | 277 | 278 | 271 | 265 |
| Minimum value | 12.8 | 17.3 | 5.81 | 5.71 |

Table 3 RFM averages

<table>
<thead>
<tr>
<th>Indicators</th>
<th>R</th>
<th>F</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average value</td>
<td>95.24248</td>
<td>1.2352</td>
<td>109</td>
</tr>
</tbody>
</table>

For the first category of users, their purchase period is less than the average, but their indicators are basically higher than the average, which means that these users are very loyal to the store and are high-value customers of the store; for the second category of users, their indicators are basically above the average except for the higher R indicator, which is relatively important to the store, but have a greater risk of churn, so define This type of user is defined as a key recovered customer. For the third category of users, all indicators are relatively average, so this category of users is defined as general customers. For the fourth category of users, their purchase cycle is greater than the average, and other indicators are lower than the average of indicators. This category of users is of little value to the store and is likely to have churned, so the fourth category of users is named churned customers. The value of each category is roughly the same as the results of the previous section based on the clustering results, with the highest value of high-value users, followed by key recovered customers, which also confirms the feasibility of applying the RFM model and K-means clustering algorithm to shoe e-commerce user segmentation and precision marketing, and the results are of some reference value. [9]

3 CONCLUSION

In recent years, with the continuous development and maturity of the e-commerce industry, more and more enterprises have started to stay in e-commerce platforms, expecting to further
expand their profits through e-commerce channels, but the entry of a large number of enterprises has led to a greater competitive pressure among e-commerce companies. In the face of a large customer base, the use of traditional marketing methods will have large investment costs, uncontrollable returns and other defects, so precision marketing has become a more reasonable marketing methods. Through big data technology to classify the store's customers, respectively calculate the value of each type of customer size, for different categories of customers to develop different marketing strategies. In this way, we can save a lot of manpower and material resources, and also improve the efficiency of marketing.

This paper establishes an RFM model for the user data of a shoe store on Taobao platform, and borrows the idea of principal component analysis to calculate the weight coefficient of each index in the RFM model when calculating the user value, and then calculates the user value according to the weight coefficient. In terms of user classification, this paper firstly adopts the clustering algorithm to classify users according to each index, and then compares the average value of different indexes of each category of users with the average value of the overall indexes to classify the customers of the store into four major categories: high-value users, key recovery users, general users and lost users. For each category of users, the total value of their users was calculated to verify the clustering results of the clustering algorithm, and the results were found to be correct and reliable.

REFERENCES

Improvement of Traditional Bond Default Identification Model Based on ESG Score

Zhiyuan Bai1,a, Nanyang Huang1,b*

baizhiyuan@stu.xjtu.edu.cn,
*Corresponding author Nanyang Huang, hny123@stu.xjtu.edu.cn

Jinhe Center for Economic Research, Xi’an Jiaotong University, Xi’an, Shaanxi 710049, China

Abstract: The most important problem in the bond default field is whether the traditional model meets theoretical expectations well. To accurately identify bond default risk, this paper establishes a theory which introduces ESG scores into the traditional model of bond default identification based on some economic theory and tests the feasibility of the theory through empirical study methods. Has there been an improvement after incorporating ESG scores? After proving that incorporating ESG can truly predict reality well and have consistent practice by using the logistic regression method, this paper uses the CatBoost method in machine learning to solve an accurate prediction model. In the end, it was found that there is a significant positive relationship between ESG and bond default risk. After analysis, this paper draws the following results. First, ESG can help us better identify bond default risk. Second, when the company judges that its operating status is not good, it will whitewash the company through earnings management and public opinion to improve its ESG score. Third, the company will also default on bonds due to improper earnings management. To sum up, the establishment of ESG theory provides investors with a more precise method to assess the risk of corporate bonds.

Keywords: Corporate Bonds, Default Rate, ESG Evaluation, Multiple Linear Regression, Stepwise Regression.

1 INTRODUCTION

Nowadays, the rapid development of China’s economic level has a certain negative effect on the environment and society. To continuously coordinate economic development with the environment and society, relevant laws and regulations have been introduced, and ESG concept has started to enter the vision of investors, enterprises, and the whole society. ESG evaluation system focuses on the performance of three non-financial indicators, namely environmental impact (E), social responsibility (S) and corporate governance structure (G), to measure the economic sustainability of enterprises.

From the existing research and practice around the world, many international organizations and investment institutions have introduced ESG concept into their original systems. At the same time, ESG has some connection with corporate credit risk. Three major credit rating agencies, Fitch, Moody’s, and S&P have announced to the public the inclusion of ESG indicators into credit rating management. As the method gradually matures and practice
advances, the application of ESG indicators on the impact of credit risk rating deserves attention.

Before 2014, China's bank wealth management products, trust wealth management products, and even fund subsidiary capital management products all had the feature of "rigid payment", which means that when the funds are at risk or the products fail to achieve the expected return, the issuer or channel party, for the purpose of interest and reputation maintenance, will guarantee the payment of the principal and return of the wealth management products by seeking third-party institutions to take over and using its own funds to advance the money first.

However, in 2014, the number of bond defaults in China exploded. In March, “11 Chaori Bond” became the first defaulted bond in the Chinese market, breaking the myth of “rigid payment”, as it could not pay the interest in full on time and reached the substantial default standard. “12 Jintai Bond” became the first bond to default on its principal. In recent years, bond default rate has gradually increased, which not only reflects the maturity of the China’s bond market, but also reveals that there is still a considerable credit risk in the market. As of 2021, there are 243 defaulting entities in China’s bond market, involving a total of 631 maturity defaulted bonds with a total maturity default amount of about 539.901 billion yuan. Analyzing the types of defaulted bonds, this paper finds that corporate bonds account for the largest proportion of the total number of defaulted bonds, which also illustrates that corporate bond credit risk deserves more attention and research.

As a bond variety with a high number of defaulted bonds, corporate bonds are representative in analyzing their specific default causes. Therefore, this paper combines ESG indicators and internal management ability with traditional corporate bond default impact indicators and constructs a system with a total of 11 variables at two levels: external to the company (macroeconomic), and internal to the company (financial indicators, ESG, and bond-related indicators), forming a more complete theoretical logic system through qualitative analysis. Moreover, with the consideration of data availability, 21 defaulted corporate bonds and 17 normal corporate maturity bonds from 2018-2021 are selected as the samples for constructing the corporate bond default risk identification model. Multiple linear regression and stepwise regression are applied to design the corporate bond default risk identification scheme, using the t-1 year data as prediction basis. Then, the results of whether ESG indicators are included are compared to determine the role of ESG information for corporate bond default risk identification.

2 LITERATURE REVIEW

Many domestic studies on ESG indicators on bond defaults and corporate ratings involve the analysis of environmental aspects, and environmental indicators should be particularly extensive in the rating of heavy industrial corporate bonds. Based on SynTao Green rating system, Zhang (2022) [1] selects credit bond issuers with defaults from 2014 to 2020 as sample data, establishes regression model analysis, and proposes that the evaluation indexes of environmental dimension mainly focus on whether enterprises have a clear environmental management system and whether they have a clear low-carbon and green plan. Using the Chinese bond ESG evaluation system as an indicator, Zhang et al. (2021) [2] analyzes ESG
situation of 52 public bond issuers that defaulted in 2017-2020 and finds that ESG scores of companies decreased year by year in the three years before default, and representative deterioration indicators in environmental subordinate indicators were the ability of companies to meet environmental requirements and prevention and control of waste gas pollution.

The social perspective in ESG covers social responsibility elements such as employee relations, supplier management, investor protection, and community and social contributions. Ruan et al. (2019) [3] shows that corporate fulfillment of social responsibility enhances capital market recognition, improves corporate reputation, and reduces financing constraints and default possibilities through various ways such as charitable donations. Zhang et al. (2021) [4] shows that corporate fulfillment of social responsibility can improve moral evaluation, reputation, and access to policy support, enhance corporate development space, and reduce the possibility of bond default in terms of the endogenous dynamics of corporate growth.

The corporate governance component of ESG indicators, as the core of sound operation and development of companies, occupies a relatively important position. Leon et al. (2020) [5] uses empirical research to show the likelihood of bond default is highly correlated with four perspectives: highly concentrated ownership, inefficient board mechanisms, low disclosure of financial indicators, and high shareholder power. Li et al. (2017) [6], on the other hand, analyzes the relationship between concurrent independent directorship and bond default risk using data from listed companies issuing bonds in the Chinese bond market between 2010 and 2013. They conclude that the board payment, as one of the cores of corporate governance, has a correlation between bond default risk and the proportion of concurrent independent directors. The greater the proportion of concurrent independent directors, the less efficient their supervision, and the greater the default risk of the company. Sun (2022) [7] constructs a bond default risk assessment index system on the correlation between corporate governance and bond default, subdivides the corporate governance risk into three secondary indicators: corporate governance competitiveness, the shareholding ratio of the largest shareholder, and the amount of funds appropriated by related parties, and checks whether the corporate governance structure and board structure are reasonable.

There are many academic approaches to evaluate bond defaults. Incomplete contract theory argues for the possibility of evaluating the default risk of a company's debt through its financial status. Debt maturity structure theory points out that the duration of debt is also an important factor affecting default. Company competitiveness evaluation theory shows that the ESG rating system can be used to make an evaluation of a company's credit risk. Rao (2020) [8] summarizes several common models, such as GBDT algorithm, AdaBoost algorithm, support vector machine, BP neural net, and CatBoost model. Combined with the findings of Chang et al. (2019) [9] that ESG scores of domestic companies drop significantly before the occurrence of credit events, and the relationship between ESG and bond defaults is well justified by using above models.

3 RESEARCH DESIGN

In recent years, bond financing, as an important means of financing in China's capital market, has grown much faster than equity financing in the same period. Since March 2014, “11 Chaori Bond” broke the “rigid payment” of China's bond market. Since then, bond defaults
have emerged one after another, and bond defaults may become “normal” in the future. Corporate bonds, as a bond with many defaults, have a certain degree of representativeness when analyzing the specific causes of default. Therefore, this paper combines ESG indicators reflecting corporate environmental responsibility, social responsibility and internal management ability with traditional corporate bond default impact indicators and constructs a system with 11 variables at two levels: external (macroeconomic) and internal (financial variables, ESG, bond-related variables), forming a more complete theoretical logic system through qualitative analysis. In addition, considering data availability, this paper selects 21 defaulted corporate bonds and 17 normal corporate maturity bonds from 2018-2021 as the samples for constructing the corporate bond default risk identification model, and applies multiple linear regression to design the corporate bond default risk identification scheme and verify the reasonableness, using t-1 year data as the prediction basis. Then, the results of whether ESG indicators are included are compared to determine the role of ESG information for corporate bond default risk identification.

3.1 Variable Selection

When selecting variables that affect corporate bond defaults, both extra-company and intra-company variables should be considered. The most significant of the extra-company variables are macroeconomic factors, while the intra-company variables cover mainly the company's finance and the bond itself.

3.1.1 Macroeconomic factors

As a result of assuming the default risk of the invested bonds, bond investors receive compensation for the risk premium, and in the bond market, the default risk of a certain bond is generally measured by comparing the difference between the interest rate of that bond and the interest rate of treasury bonds (risk-free rate) in the same period, which is also known as the credit spread. Since the macroeconomic development directly affects the Treasury rate, the credit spread associated with the Treasury rate is also directly affected, which in turn affects the company's financing costs. The indicator of annual GDP growth rate fully reflects the current macroeconomy. When annual GDP growth rate slows down or even becomes negative and the macroeconomy is in a downward cycle, investors tend to increase the weight of the Treasury rate in their portfolios, and companies can only attract investors by raising the interest rate of corporate bonds, which leads to high financing costs and can easily cause a break in the company's capital chain, which in turn increases the risk of default.

3.1.2 Company Financial Factors

Financial information is the most direct information to judge the company's operating condition, and investors can judge the company's solvency, profitability, operating capacity, and growth capacity from the published financial information.

Specifically, solvency indicators visually reflect a company's ability to repay its debts and indicate whether the company has sufficient cash flow to repay its short-term and long-term debts. If a company has a low solvency ratio, it is more likely to default on its debts. Profitability indicators visually reflect a company's ability to generate earnings, and are representative of return on assets, return on net assets and earnings per share. If a company’s earnings are volatile and unstable, it is likely to default on its bonds. Operating capacity
indicators visually reflect a company's ability to operate its assets, and a representative indicator is accounting receivable turnover. If a company is slow to recycle funds and its working capital flow is easily disturbed by various external factors and thus unstable, the company has a high risk of default. The growth ability indicator intuitively reflects the company's potential for sustainable operation, and the representative indicator is the growth rate of operating income. If a company's operating income is stagnant and business growth is slow, the company has a higher risk of default.

3.1.3 Factors of bond itself

The indicators related to the bonds themselves also reflect the risk of bond default to a certain extent. The higher the coupon rate of the bond and the larger the actual issue volume, the greater the pressure on the company's operating cash, the greater the pressure to repay the bond, and thus the higher the possibility of bond default.

3.2 Data Introduction

This paper collates all the bonds issued by enterprises with bond defaults in China's real estate industry from 2018-2021 and obtains 38 relevant bonds. Since there are many industries involved in bond defaults and large differences between different industries, it is difficult to compare them, thus this paper selects real estate enterprises with defaults after 2018, and a total of four enterprises with bond defaults are selected from Huaxia Happiness Foundation Co, Beijing Huaye Capital Holdings Co, Zhonghong Holdings Co, and Taihe Group Co. The data on the firm characteristics variables in this paper are obtained from the Guotaian database and Bloomberg.

3.3 Data processing

3.3.1 Assign values to special data

Examining the data, this paper finds that bonds are classified according to either “default” or “normal” status. To facilitate the data processing, if the bond status is "default", it is assigned a value of 1, which means the default probability is 100%; if the bond status is "normal", it is assigned a value of 0, which means the default probability is 0%.

3.3.2 Processing of missing values

By checking the data, this paper finds that there is no ESG rating for Zhonghong Holdings Co. in ESG rating data queried by Bloomberg, and randomly checking ESG ratings given by Shang Dao Rong Green, this paper finds that Zhonghong Holdings Co. and Taihe Group Co. have the same.

<table>
<thead>
<tr>
<th>Table 1: Definition of Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type</strong></td>
</tr>
<tr>
<td>Dependent Variable</td>
</tr>
<tr>
<td>Independent Variable</td>
</tr>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>VIF</td>
</tr>
</tbody>
</table>

3.4 Model Construction

3.4.1 Variable definition

There are one dependent variable and eleven independent variables in the regression model in this paper. Please see table 1.

3.4.2 Model building

This paper explains the effect of ESG on bond default probability by constructing the following multiple linear regression model without ESG factors.

\[
Y = \alpha_1 X_1 + \alpha_2 X_2 + \beta_1 alr + \beta_2 eps + \beta_3 roa + \beta_4 roe + \beta_5 oig + \beta_6 tar + \gamma_1 gdp + \varepsilon
\]

The model with ESG is as follows.

\[
Y = \alpha_1 X_1 + \alpha_2 X_2 + \beta_1 alr + \beta_2 eps + \beta_3 roa + \beta_4 roe + \beta_5 oig + \beta_6 tar + \gamma_1 gdp + \varnothing_1 esg + \varepsilon
\]

ESG rating in 2018, and had similar ratings in the E, S, and G subscales, so it can be inferred that these two companies had similar ESG status in 2018, so the ESG rating of Taihe Group Co. in Bloomberg in 2018 was used instead of Zhonghong Holdings Co. for analysis.
4 EMPIRICAL ANALYSIS

4.1 Multiple linear regression analysis

Multiple linear regression is a statistical analysis method that examines the linear relationship between a continuous response variable and multiple explanatory variables. Like simple linear regression, multiple linear regression examines the regression coefficients, the R-squared, the test and the conditions for holding.

Regression analysis was first done on all independent variables without the ESG score model and default rates, with the following results:

<table>
<thead>
<tr>
<th>Table 2: VIF of Each Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-squared</td>
</tr>
<tr>
<td>Mean VIF</td>
</tr>
<tr>
<td>White test</td>
</tr>
</tbody>
</table>

The table 2 shows that although the fit has improved slightly after the inclusion of the ESG scores, the VIF values have also increased significantly compared to before the inclusion, and several variables still have VIF values greater than 10, and there is multicollinearity, which still does not satisfy the conditions for multiple linear regression analysis.

Therefore, with or without the inclusion of ESG scores as a factor, trying to use multiple linear regression to explore the relationship between the dependent and independent variables requires trying new methods for improvement.

4.2 Stepwise regression-based model improvement

As the variables did not satisfy the conditions for multiple linear regression, this paper used stepwise regression to remove variables with co-linearity. The basic idea is to reduce the degree of multicollinearity by removing variables that are less important and highly correlated with other variables. The explanatory variables other than ESG scores were first regressed stepwise against the default rate, with the results in table 3.

<table>
<thead>
<tr>
<th>Table 3: Regression Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non standardized coefficient</td>
</tr>
<tr>
<td>B</td>
</tr>
<tr>
<td>Constant</td>
</tr>
<tr>
<td>EPS</td>
</tr>
</tbody>
</table>
The actual issue volume, coupon rate, gearing ratio, return on net assets, earnings per share, return on assets (ROA), operating income growth rate, accounts receivable turnover rate and annual GDP growth rate were used as independent variables, while the default rate was used as the dependent variable in the stepwise regression analysis. And the $R^2 = 0.306$, which can explain 30.6% of the variation in the default rate. Furthermore, the model passed the F-test ($F=15.873, p=0.000<0.05$), indicating that the model is valid.

To test the normality of earnings per share, the P-P diagram test is then applied, which reflects the extent to which the actual cumulative probability of the variable matches the theoretical cumulative probability and can be used to examine whether the data obeys a certain type of distribution. If the data obeys a normal distribution, the data points should largely coincide with the theoretical straight line. As shown, the image is found to be approximately diagonal and the data points largely coincide with the theoretical straight line, indicating some normality.

The model formula is:
Default rate = $0.782 - 0.178 \times \text{eps}$

The regression coefficient value for earnings per share is $-0.178 \ (t=-3.984, \ p=0.000<0.01)$, implying that earnings per share can have a significant negative relationship on the default rate.

The heat map plots the correlation coefficients between the explanatory variables without ESG, with darker red indicating a stronger positive correlation and darker blue indicating a stronger negative correlation (fig. 2).
The higher calorific values in the upper left and middle regions were observed to be more relevant.

Through the above analysis, the performance of bond defaults over the period 2018-2021 does not fit with bond default theory through empirical analysis tests, this paper next proceeds to add ESG to the model for testing. The results are as follows:

<table>
<thead>
<tr>
<th>Indicators</th>
<th>Regression coefficients</th>
<th>95% CI</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-2.574** (-8.554)</td>
<td>-3.164--1.985</td>
<td>-</td>
</tr>
<tr>
<td>EPS</td>
<td>-0.415** (-10.722)</td>
<td>-0.491--0.339</td>
<td>4.24</td>
</tr>
<tr>
<td>ROE</td>
<td>2.316** (8.948)</td>
<td>1.809--2.823</td>
<td>5.13</td>
</tr>
<tr>
<td>OIG</td>
<td>-0.315** (-11.331)</td>
<td>-0.370--0.261</td>
<td>3.40</td>
</tr>
<tr>
<td>TAR</td>
<td>0.020** (4.195)</td>
<td>0.010--0.029</td>
<td>1.85</td>
</tr>
<tr>
<td>ESG</td>
<td>0.189** (11.771)</td>
<td>0.158--0.221</td>
<td>4.22</td>
</tr>
<tr>
<td>Sample size</td>
<td>38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R²</td>
<td>0.891</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.874</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F value</td>
<td>F(5,32)=52.441,p=0.000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The actual issue volume, coupon rate, Balance sheet ratio, eps, roe, Operating income growth rate, accounts receivable turnover, GDP growth rate, esg as independent variables, and default rate as dependent variable were used in the stepwise regression analysis, and after automatic identification by the model, the remaining NAV, EPS, operating income growth rate, accounts receivable turnover, and ESG were identified. A total of five items were included in the model, with an R-squared value of 0.891, implying that ROE, EPS, operating income growth rate, accounts receivable turnover rate and ESG could explain 89.1% of the variation in the default rate. Moreover, the model passed the F-test (F=52.441, p=0.000<0.05), indicating that the model is valid. The normality of earnings per share was then tested by a P-P diagram (as shown), and the images of the selected independent variables were found to be approximately diagonal, indicating some normality.

The model equation is:

Default rate = -2.574 + 2.316*roe - 0.415*eps - 0.315*oig + 0.020*tar + 0.189*ESG

In addition, a retrospective test for multicollinearity in the model revealed that the VIF values of several variables in the model were greater than 5 but less than 10, implying that there may be some cointegration problems, so the closely correlated independent variables were checked and the analysis was re-run after eliminating the closely correlated independent variables. The final specific analysis revealed that:
(a) The value of the regression coefficient of ROE is 2.316 ($t=8.948$, $p=0.000<0.01$), implying that NPA will have a significant positive relationship on the default rate.

(b) The regression coefficient value for earnings per share is -0.415 ($t=-10.722$, $p=0.000<0.01$), implying that earnings per share can have a significant negative relationship on the default rate.

(c) The regression coefficient value for Operating income growth rate is -0.315 ($t=-11.331$, $p=0.000<0.01$), implying that the growth rate of operating income will have a significant negative relationship on the default rate.

(d) The regression coefficient value for accounts receivable turnover is 0.020 ($t=4.195$, $p=0.000<0.01$), implying that accounts receivable turnover will have a significant positive relationship on default rate.

The value of the regression coefficient for ESG is 0.189 ($t=11.771$, $p=0.000<0.01$), implying that ESG will have a significant positive relationship on the default rate.

A heat map of the correlation coefficients between the explanatory variables for the inclusion of ESG is plotted in fig.3.

![Figure 3: Heat map of correlation coefficients between explanatory variables including ESG](image)

The heat map of correlation coefficients between explanatory variables relative to those that do not include ESG is less red as a proportion of the full plot, indicating that it has fewer multicollinearity variables than those that do not include ESG, consistent with the results of the stepwise regression method (only one variable remains after stepwise regression).

Thus, by incorporating ESG into the model, more of the independent variables pass the stepwise regression test, including the ESG score, which is partially in line with theoretical expectations.

However, as the bond default rate has a value of 0 or 1, like a dummy variable, it would be more natural to treat it as a dummy variable.
4.3 Model improvement based on machine learning CatBoost

The use of multiple linear regression cannot handle bond default models that incorporate or exclude ESG, and stepwise regression, while able to explain the effect of some variables on bond defaults after reducing some of them, still excludes factors that should theoretically be included. Now this paper uses CatBoost algorithm to construct a corporate bond default risk identification model as an improvement of the regression model, and thus designs a corporate bond default risk identification scheme incorporating ESG information based on the CatBoost algorithm.

Without considering ESG:

The parameters were first designed, and the specific values were selected as shown in the following table:

<table>
<thead>
<tr>
<th>Table 5: Parameter Design Value Selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Slicing</td>
</tr>
<tr>
<td>Data Shuffle</td>
</tr>
<tr>
<td>Cross-validation</td>
</tr>
<tr>
<td>Number of iterations</td>
</tr>
<tr>
<td>Learning Rate</td>
</tr>
<tr>
<td>L2 canonical term</td>
</tr>
<tr>
<td>Maximum depth of the tree</td>
</tr>
<tr>
<td>Overfitting detection threshold</td>
</tr>
<tr>
<td>Number of iterations to continue after reaching optimization</td>
</tr>
</tbody>
</table>

The training was carried out after the parameter settings were made and the following results were obtained:

<table>
<thead>
<tr>
<th>Table 6: Results of Training Set and Test Set</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSE</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>Training set</td>
</tr>
<tr>
<td>Test set</td>
</tr>
</tbody>
</table>

In summary, using the Catboost method can predict a sample fit of 0.926, which is a significant improvement over the previous one, but with a larger MAPE value.

A graph of the test data predictions is shown in fig.4, which shows a good trend and fit.
Now consider the case where ESG scores are integrated into the model. The parameter design was the same as before. After performing the training with the parameter settings, the following results were obtained.

<table>
<thead>
<tr>
<th>Table 7: Results of Training Set and Test Set</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSE</td>
</tr>
<tr>
<td>-----</td>
</tr>
<tr>
<td>Training set</td>
</tr>
<tr>
<td>Test set</td>
</tr>
</tbody>
</table>

Seeing a significant reduction in the MAPE value from before, while the fit is 0.931, another improvement from before, indicating a more accurate evaluation of bond defaults after incorporating the ESG score.

A graph of the test data predictions is shown below, which shows a good trend and fit.
4.4 Improvement of CatBoost prediction model based on CatBoost classification

In line with the idea of using logit regression to improve multiple linear regression models, CatBoost regression is suitable for making predictions on continuous variables, but the default rates used in this paper are discrete variables, so this paper use the CatBoost classification model instead of the regression model.

This paper used the Catboost classification for each of the two cases with or without ESG, with the same parameter settings as in the Catboost regression, to produce the following results:

Figure 6: A graph of the test data prediction (ESG scores are integrated)

Figure 7: The importance of all features (ESG scores are integrated)

Figure 8: The importance of all features
5 CONCLUSION

By comparing the two models in the linear regression and whether to include the variable ESG score, the precision of the judgement of the default rate of corporate bonds can be improved through the ESG score. Since there is a positive relationship between the ESG score of a company's T-1 year and the default rate of its bonds, which is not in line with previous expectations, this paper speculates that: firstly, companies improve their ESG scores through surplus management as well as through public opinion and whitewashing when they judge themselves to be in a poor state of operation; secondly, defaults on corporate bonds in the real estate sector are caused by prior mismanagement of surplus. Furthermore, the CatBoost approach still leads us to the same conclusion as the linear regression, namely that the incorporation of ESG will help companies to discriminate between bond default risk, but whether ESG has a positive or negative impact on this is more difficult to determine through the importance of the characteristics of ESG scores. This paper also finds that the inclusion of ESG in the explanatory variables is effective in reducing the multicollinearity between variables.

REFERENCES

Research on Innovative Practice of Financial Management of Power Grid Enterprises Based on RPA Technology

Changping Chen\textsuperscript{1,a*}, Yansheng Li\textsuperscript{1,b}, Wenjie Yang\textsuperscript{1,c}, Bing Lin\textsuperscript{1,d}

\*Corresponding e-mail: chenchangping@sgitg.sgcc.com.cn, liyansheng@sgitg.sgcc.com.cn, 358625210@qq.com, linbing@sgitg.sgcc.com.cn

AOSTAR INFORMATION TECHNOLOGIES CO., LTD. Chengdu 610000, China

Abstract: With the remarkable progress of modern science and technology, the application of a series of new technologies, such as big data, information technology, science and technology smart city, has been expanding, making people's daily life more convenient. In the financial management of power and digital companies, transformation should be achieved. The technology of RPA, such as automated software tools, has been widely used in the field of financial management. It can realize intelligent management through application scenarios, improve automation level and financial management efficiency, and reduce the financial risk of enterprise power. Therefore, based on the technology of RPA, this paper proposes to analyze and discuss the digital transformation of financial management of electric power enterprises, hoping that it will help improve the financial management level of electric power enterprises.

Keywords: RPA, Electric Power Enterprises, Financial Management, Innovate.

1 INTRODUCTION

Under the trend of economic globalization, China's economic development is gradually moving towards digital and intelligent reform, and various new intelligent technologies are used to improve the speed of economic and social development. It is necessary for power companies to promote the digital transformation of relevant service fields in the new era \cite{1}. Among them, financial management is one of the basic departments of power companies \cite{2}. In order to avoid too much duplication. The basic work in the work process is that RPA technology can be used to realize the digital transformation of the financial management of power companies. Therefore, this paper analyzes the digital financial management of energy conversion. Electric company based on RPA technology.

2 OPTIMIZATION OF FINANCIAL REIMBURSEMENT OF POWER ENTERPRISES BASED ON RPA TECHNOLOGY

2.1 Digital Management of Financial Reimbursement of Electric Power Enterprises

The reimbursement process of power enterprises optimized by RPA technology is shown in
Figure 1.

Electric power enterprises use accounting software to directly extract relevant information from electronic invoices, and generate accounting vouchers and statements directly, which avoids mistakes that often occur in the process of manual operation. Promote the use of electronic invoices to accelerate the development of financial integration, digitalization, automation and informatization of power enterprises, and solve the problems related to financial management from the source [3].

Figure 1 Reimbursement Process of Power Enterprises Optimized by RPA Technology
2.2 Reimbursement Violation Prediction Based on Decision Tree

Suppose that the preprocessed data is the training data set G, a series of tests are carried out on the reimbursement related content, and the number of iterations is set, that is, all scores can reach the leaf node and the decision tree results can be entered. The whole operation process is shown in Figure 2.

The process is described as follows:

1. The preprocessed violation prediction data object is set as the data set G, X Property set for violation or not:

   \[ G = \{ g_m | g_m = (g_{m1}, g_{m2}, ..., g_{mn}) , m = (1,2,...,k) \} \]  

   Where, \( g_m \) represent the data set \( m \) Data, \( n \) indicates that each piece of data has \( n \) Attributes, that is, the input data attributes of the violation data set \( X \). It includes two target attributes to determine whether it is a violation. Violation and non-violation are represented by 1 and 0 respectively.

2. Test condition division. The degree of uncertainty and confusion of information is described by information entropy. Set sample set \( G \) Memory in \( m \) Class samples, and the probabilities are \( p_k (k = 1,2,...,m) \). Define the set using equation (2) \( G \) Information entropy:
\[
\text{Ent}(G) = \sum_{k=1}^{m} p_k \log_2 \frac{1}{p_k} = - \sum_{k=1}^{m} p_k \log_2 p_k \quad (2)
\]

Where, \( 0 \log_2 0 \) Defined as 0, at this time \( \text{Ent}(G) \) The higher the value of, the higher the purity of the set G, which leads to higher uncertainty and confusion of information. If the partition condition is variable C, partition the set \( G \) by \( n \) Subsets, use formula (3) to represent each subset \( G_i \). Information entropy of:

\[
\text{Ent}(G|C) = \sum_{i=1}^{n} \frac{N(G_i)}{N} \text{Ent}(G_i) \quad (3)
\]

Where, \( N \) And \( n \) They respectively represent the number of samples in the parent node and the number of test condition groups, \( N(G_i) \) Indicates the sample quantity of each grouping subset.

The purpose of the information gain is to compare the purity difference between the child node and the parent node, and measure the effect of the test conditions. If the test conditions are good, large differences are required. Equation (4) represents the information gain:

\[
\text{Gain}(G,C) = \text{Ent}(G) - \text{Ent}(G|C) = \text{Ent}(G) - \sum_{i=1}^{n} \frac{N(G_i)}{N} \text{Ent}(G_i) \quad (4)
\]

The information gain is used to partition. The small information entropy is due to the large number of subsets. The purity of the subset is reduced by merging multiple classification attributes, that is, the secondary classification. The gain rate evaluation condition of the decision tree algorithm is used to avoid excessive classification. The result of the combined partition is as follows:

\[
\text{Gain}_{\text{ratio}}(G,C) = \frac{\text{Gain}(G,C)}{\text{Ent}'(C)} \quad (5)
\]

\[
\text{Ent}'(C) = \sum_{j=1}^{k} \frac{N(G_j)}{N} \log_2 \frac{N(G_j)}{N} \quad (6)
\]

Use of deviation caused by multiple partitions \( \text{Ent}'(C) \) Correction, that is, the cost of selecting partition attribute C.
A decision tree with high accuracy is obtained by continuously dividing the original data. The decision tree is pruned to reduce over fitting. The pruned decision tree is used to predict violations of employee reimbursement data.

3 RESEARCH ON THE APPLICATION OF RPA TECHNOLOGY IN THE INNOVATIVE PRACTICE OF FINANCIAL MANAGEMENT OF POWER GRID ENTERPRISES

3.1 Application of daily work of financial management

3.1.1 Application of settlement professional pilot

In order to avoid the capital payment risk of power grid enterprises, RPA technology is applied to the settlement professional pilot. The advanced "payment monitoring process interception tracking and notification instruction" function can automatically intercept funds according to the rules specified in the payment process, and then send the relevant information to the financial personnel regularly. After receiving the system information, the financial personnel constantly update the payment system to check in real time for interception. In the process of updating the information, the RPA media can automatically filter the instructions that have not responded for a period of time very long. At this time, the system will send the relevant notice to the customer according to the set contact list of the payer to ensure the smooth completion of the fund settlement business [4].

3.1.2 Application of invoice verification pilot

In the new era, in order to meet the challenge of the extensive development of the electric ticket business, the "industrial chain invoice verification" function developed by RPA technology can automatically verify invoices, so as to ensure the effectiveness of the electric ticket business, and automatically enter the invoice information after inspection into the system database. Before technical optimization, financial personnel need to constantly refresh invoice query information to judge whether the data is effectively used. This link is time-consuming and inefficient. After using RPA technology, the robot automatically logs into the company's grid financial management system, and records the inspection results in the form of transmission according to the relevant provisions of invoice management and financial management methods. At the same time, the robot will automatically complete Log in to the VAT invoice verification platform of the State Administration of Taxation, and automatically generate the corresponding electronic files for the inspected invoices.

3.1.3 Application of customer service pilot

In order to further improve the utilization efficiency of data and information resources of resources and ensure the accuracy of customer information, RPA technology is applied to optimize the original financial management work. The "customer information screening and verification" function developed can automatically compare customer information, establish good contact with the industrial and commercial registration system externally, test the internal customer information of the system, and promote the continuous improvement of
Under the traditional financial management mode, employees need to manually compare customer data of complex types and large quantities, which inevitably leads to data omissions, errors and other problems. After the application of RPA technology, the robot can lock customers in advance according to the data stored in the system, and collect customer data according to the data information comparison results, which not only improves customer experience and satisfaction, but also fundamentally solves the problems of inaccurate manual verification, untimely information update, and incomplete information collection, Lay a solid foundation for the digital transformation of enterprise power grid financial management.

3.1.4 Application of data statistics pilot

Under the original financial management mode, power grid enterprises need special financial personnel to integrate and process reports in different financial systems, Then the comprehensive results will be transmitted to different departments to support administrative decision-making. In order to further improve the data quality and the efficiency of providing information, the RPA has introduced the function of "automatically generating statistical reports", which can be linked. The validity of data from different systems, replace manual operation with automatic generation of reports, and promote the improvement of business processing efficiency and financial management information level of power grid enterprises[5].

3.2 Risk Control of Application

3.2.1 Operational risk

The operational risk of RPA technology application is closely related to the operation of financial personnel. In order to avoid the occurrence of operational risks, on the one hand, targeted training will be carried out to strengthen the ability of financial personnel to master and apply the standardized process; on the other hand, at the beginning of designing RPA, each standardized scenario will be reasonably set, so as to delay the business black box trend of long-term application of RPA, achieve stable and smooth financial operation procedures, and ensure the digital transformation of financial management of power grid enterprises.

3.2.2 System risk

RPA robot usually works in coordination with the original financial system and server. If the system updates and upgrades abnormally or the system is incompatible, RPA will have the problem of program termination in serious cases [6]. If RPA is unable to automatically adjust the abnormal response in a short time or stops running suddenly, it will have an unpredictable impact on the financial management of the enterprise. Therefore, it is necessary to take timely prevention and control measures to avoid system risks. System users should pay attention to the preservation of daily financial management records and logs, and do a good job of data backup, so as to avoid the sudden interruption of RPA and affect the smooth development of enterprise business.

3.2.3 Personnel management risk

The application of RPA technology means that the number of posts involved in repetitive operations in power grid enterprises will decrease, which will change the labor structure of financial management in enterprises. In order to give better play to the advantages of RPA
technology application, enterprises should formulate corresponding financial personnel learning plans for RPA technology application goals, and establish an IT department that coordinates with the financial department to avoid business operation risks caused by adverse team operations.

3.2.4 Information security risks

Personal information and confidential information are important components of the financial data information of power grid enterprises. The financial management system that uses RPA technology innovation will be affected by malicious access and wrong use, resulting in the risk of information leakage. In response to such risks, enterprises should set up a strong information security inspection organization and effectively improve the information security protection mechanism to deal with information security risks. When executing the RPA process, the system will program the ID and password into the files to be processed by the terminal in advance to avoid the risk of information theft caused by malicious operations.

3.3 Technical Application Benefits

3.3.1 Management benefits

As the power grid enterprises shoulder more social responsibilities and obligations, the application of RPA technology is of great significance for enterprises to improve business processing efficiency and resource utilization efficiency. After deeply combining RPA technology with artificial intelligence technology, we will develop an artificial interactive intelligent robot to automatically manage enterprise material data, which will help further improve the enterprise material data management system. The intelligent robot can fully perceive the situation and efficiently use RPA technology to handle the operation business of power grid enterprises, so as to shift the focus of enterprise work to a new management mode that uses data for analysis and decision-making, and realize the improvement of traditional management mode and system process.

3.3.2 Economic benefits

After the introduction and application of RPA technology in enterprise financial management, it can complete a variety of types and a large number of tasks in a short time, and can operate all day without interruption and negative emotions, so as to meet the needs of power grid enterprises for multi platform processing of financial management. It has high stability and greatly saves the operating costs of various resources of the enterprise [7]. Because the business systems of power grid enterprises have different caliber, the application of RPA technology has effectively solved the problem of industry finance integration, effectively reduced the cost of later integration and transformation, and achieved the goal of connecting industry finance data of all businesses and systems.
4 APPLICATION OF DIGITAL OPERATION MODE FOR FINANCIAL MANAGEMENT OF POWER GRID ENTERPRISES

4.1 Business Demand Analysis And Assessment

Based on the process maturity evaluation, frame the application scope of the company's financial process automation, form a priority framework for the application of financial process automation and an expansion path for industrial and financial process automation, define the application scope, and form a three-level standard. First, financial foundation, unifying process and data standards, forming financial process automation, and giving priority to implementation; The second is operation upgrading. After the process is automatically completed, the operation of key links will be upgraded to deepen the integration of industry and finance; The third is to support decision-making, business planning and analysis, decision-making and support.

4.2 Application Scenario Practice

4.2.1 Basic Financial Management - General Ledger Management

In order to ensure the accuracy and rationality of the group's accounting, ensure the basic financial data, liberate financial personnel, and improve the efficiency of industry finance integration, the company needs to collect the engineering projects, entry accounts, entry amount, cash flow and other information of each unit one by one in the system according to the "consolidated cross tier company list", which takes a long time, and may lead to collection errors due to the large amount of data. The rule robot is used to clarify the rules based on the three table data of "project settlement", "product purchase, labor service acceptance" and "cash flow", and automatically complete the cross secondary unit consolidation and offset data collection, which has improved the efficiency by nearly 9 times (see Figure 3).

![Figure3 Data Collection of Cross secondary Unit Consolidation and Offset](image-url)
4.2.2 Basic financial management - fixed assets accounting

For power grid enterprises with large assets and many projects under construction, in the process of completing project asset transfer and carrying out fixed asset accounting, rule robots are used to establish rules around data query, amount verification, settlement rule maintenance, asset card creation and other scenarios, effectively solving the process work of multiple data types, large data scale, high repeatability, and long time consumption, and ensuring the consistency of fixed asset information. The workload of 18 days per month can be replaced by labor (see Figure 4).

![Figure 4 Project transfer](image)

4.2.3 Financial intelligent management - industry finance integration analysis

In order to promote the deep integration of industry and finance, further advance the management of financial expertise [10]. Taking photovoltaic new energy settlement as an example, business personnel need to have a lot of professional financial knowledge in the process of processing, review relevant business information manually, coordinate business personnel to solve problems after finding them, use knowledge robots to complete problem business identification, and rule machine people to complete automatic adjustment of errors, effectively solving the business scenarios with high repeatability in the settlement and payment stages. The accuracy and efficiency of work have been greatly improved (see Figure 5).
In the data verification stage, the rule robot is used to automatically export the system data if there is a difference in the data according to the predetermined judgment logic, compare it with the original EXCEL table line by line, find out the difference lines and mark them, and notify the business personnel for further processing through the interactive robot. If the data is correct, the rule robot will analyze the difference between the actual amount of the project and the cash flow plan, and the rule robot will summarize and generate a settlement document, which will be automatically transferred to the approval post to generate settlement vouchers.

In the payment processing stage, the rule robot cooperates with the interactive robot to initiate payment, generate a payment application form, and generate a prefabricated payment voucher for the settled documents, which are automatically transferred to the review post for review and posting.

5 SUMMARY

To sum up, the global acceleration has brought new opportunities and challenges to the operation and management of power grid enterprises. In order to avoid the shortage of financial management and increase the probability of enterprises confronting risks, It is urgent to combine the technology of the RPA with the internal financial control, standardize the financial information of the RPA, improve the financial management system, improve the quality of personnel, and achieve The goal of improving financial management automation, It provides continuous power for the digital transformation of financial management of power grid enterprises.

REFERENCES

Economic Development Prediction Model Based on Deep Convolutional Neural Network

Taisong Liu1*, Xiaona Cai1, Songyu Xie2, Shanting Tan3, Zihan Ying4

*Corresponding Author E-mail Address: 2224628031qq.com

College of Economics and Trade, University of Electronic Science and Technology of China, Zhongshan Institute, Zhongshan City, China1
College of Management, Guangdong Peizheng College, Guangzhou City, China2
School of Computing, University of Electronic Science and Technology of China, Zhongshan Institute, Zhongshan City, China3
School of Foreign Languages, Guangdong Technology College, Zhaoqing City, China4

Abstract: Economic prediction is an essential method for arranging the development strategies and provide reliable data for the managers. However, previous researches about prediction were primary concentrated on the mathematical model by utilizing various statistic or economic theories, which are not considering the real situation parameters including cultural affects, political aspects and real-world factors. In this article, we utilize deep convolutional neural network to train a neural model by utilizing the ten years from 2010 to 2020 economic development parameters and predict the 2021-2022 economic development results in GuangZhou city of China. In our proposed model, the model is consisted by three primary components including pro-processing selector by utilizing the normalization, multiple deep convolutional network and multilayer perceptron to provide the final prediction results. From our extensive experimental steps, we can observe that our proposed mechanism can precisely provide the development tendency in 2021 with acceptable computation cost.

Keywords: Economic prediction, Deep convolutional neural network, Normalization, Multilayer perceptron, Computation cost.

1 INTRODUCTION

Economic prediction is an issue about speculation and estimation of future scenarios of economic phenomena. Researcher [1] based on the history and current situation of the economic development process. Scientific forecasting methods are used to reveal the development law of economic phenomena and the corresponding interconnection parameters between various economic phenomena [2-4] and point out the future development trend and possible level of economic phenomena is proposed by researchers [5]. The content of economic forecasting is very broad. Initially, the forecast of the domestic economic situation, such as the development trend of production, growth rate, economic structure, price change trend, population employment, changes in fiscal revenue and expenditure, and the supply, production and sales of various products is proposed by researchers [6-8].
At the same time, it is necessary to predict the international economic situation, such as international economic fluctuations and changes in the international market. However, extensive methods were proposed to dispose the economic prediction problems including establishing statistical analysis model, utilizing economic theories construct prediction process, training machine learning to quantify the economic inductors and predictions modelling methods is established by researchers [9-15]. In this paper, we propose a novel deep convolutional construction to predict the economic development tendency and simulate the model with the real economic data to predict the two years of developments.

Economic impacts are mainly divided into direct economic impacts and indirect impacts. The direct economic impact involves industries such as retail sales, tourism, and integrated services. Many aspects are difficult to quantitatively evaluate, and this paper uses deep convolutional neural network models to quantitatively evaluate the impact of retail, tourism, and integrated services.

Following includes the structures of this article. In section 2, we will introduce the basic knowledge of related our proposed method and primary symbols that used in our paper. Section 3 will demonstrate the main framework of our model and section 4 illustrates the experimental results that is simulated with current economic prediction models. Finally, we will conclude our main contributions and provide the future improvements methods in section 5.

2 PRELIMINARY AND SYMBOL DESCRIPTION

In this section, we will introduce the basic theories of deep convolutional neural network and the principles that used in our model. Additionally, we also illustrate the primary symbols and its descriptions in following Table 1.

2.1 Deep Convolutional Neural Network

Deep convolutional neural networks are mainly composed of input layer, convolutional layer, activation function, pooling layer, fully connected layer and output layer. The deep convolutional network can directly take the image as the input of the network, extract features through training, convolution operation, convolution operation and pooling operation, and perform operation output through the fully connected layer.

The convolutional neural network mimics the visual perception mechanism of living organisms, which can carry out supervised learning and unsupervised learning, and its implied convolution kernel parameter sharing within the layer and the sparsity of the interlayer connection make the convolutional neural network can be used to lattice the features with a small amount of computation.

Convolutional neural network is a kind of feedforward neural network with deep structure that contains convolutional computation, and is one of the representative algorithms of deep learning. Convolutional neural networks have the ability of representation learning, and can classify input information by translation invariant according to their hierarchical structure, which is a translational invariant artificial neural network. The convolutional layer simulates
human visual perception, that is, local perception function. Convolutional layers are a way of extracting data features.

2.2 Primary Parameter Symbols and Description

Following table contains the primary parameters and corresponding descriptions that were used in this article.

<table>
<thead>
<tr>
<th>Parameter Symbols</th>
<th>Explanations</th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
<td>Input samples</td>
</tr>
<tr>
<td>x'</td>
<td>Effective samples</td>
</tr>
<tr>
<td>L(x)</td>
<td>Standard division operation</td>
</tr>
<tr>
<td>N</td>
<td>Number of samples</td>
</tr>
<tr>
<td>Conv</td>
<td>Convolutional operation</td>
</tr>
<tr>
<td>MAX</td>
<td>Selection maximum function</td>
</tr>
<tr>
<td>ATT</td>
<td>Attention operation function</td>
</tr>
<tr>
<td>W</td>
<td>Trained weighted value</td>
</tr>
</tbody>
</table>

3 SYSTEM FRAMEWORKS

In this chapter, we specifically analysis the three components and show the framework in each sub-section.

3.1 Data Pro-processing Stage

Initially, the pro-processing is generating the normalized data and select the effective data for the deep neural network. The general framework of pro-processing procedure is demonstrating in following Figure 1.

Figure 1: Pro-process stage framework.
In order to eliminate the effect when different attributes of the sample have different magnitude, the difference of orders of magnitude will lead to the dominance of attributes of larger magnitude, and the difference of orders of magnitude will cause the speed of iterative convergence to slow down, and algorithms that depend on sample distance are very sensitive to the order of magnitude of the data. Normalize data based on the mean and standard deviation of the original data. Normalize the original value $x$ to $x'$ using z-score. The z-score normalization method is suitable for situations where the maximum and minimum values of the attribute are unknown, or when there is outlier data that is out of the value range. Generated data is equal to original data minus mean divided standard deviation. Equation (1) demonstrates the detail process of data selection.

$$x' = \sum_{i=0}^{N} (x - \bar{x})/L(x)$$

(1)

3.2 Framework of Deep Convolutional Neural Network

In this sub-section, we detail demonstrate the multiple layers of our proposed deep convolutional neural network structure and explain the detail function of each layer.

![Figure 2: Framework of Proposed Deep Convolutional Neural Network.](image)
In above Figure 2 demonstrates the general structure of proposed deep convolutional neural network including multiple convolutional layers represents as \texttt{Conv}, activation function is utilizing linear activation represents as \texttt{ReLU} layer and softmax layer is working as fully-connected with trained weighted edges for each connection.

Additionally, from above figure the utilization of aggregation operation is representing in following Equation (2).

\[
agg(a, b) = \frac{1}{2I} \sum_{j=0}^{I} MAX(a_j, b_j) \cdot S(a_j, b_j) \tag{2}
\]

### 3.3 Multi-layer Perceptron

The multilayer perceptron is composed of input and output layers and there can be multiple hidden layers in the middle and the prediction information is gradually transmitted from the first layer to the higher level.

A multilayer perceptron is a forward-structured artificial neural network that maps a set of input vectors to a set of output vectors. A multilayer perceptron can be thought of as a directed graph consisting of multiple node layers, each connected to the next. In addition to the input node, each node is a neuronal with a nonlinear activation function called a processing unit. A supervised learning method called the backpropagation algorithm is often used to train multilayer perceptron. Multilayer perceptron is widely used the principles of the human nervous system, learn and make data predictions.

It first learns, then stores the data using weights, and uses algorithms to adjust weights and reduce bias during training, for specifically, errors between actual and predicted values. The main advantage is its ability to solve complex problems quickly.

The parameters of the multilayer perceptron are the connection weights and biases between the layers and all parameters are randomly initialized. After iteratively trained, the gradient and update the parameters are continuously calculated until a certain prediction accuracy is greater than the expected value. Figure 3 demonstrates the detail structure of multi-layer perceptron (MLP) structure and the weighted value is representing as \( W \).
4 EXPERIMENTAL RESULT AND ANALYSIS

In this section, we simulate our proposed model to predict the economic development of city Guang Zhou in China and compare our result with other existing prediction models named RBF neural network is proposed by (Bin Li and others, 2022) [16] and normal GM Grey Method is established by (Ying Yu and others, 2022) [17] respectively. From our extensively experimental result and analysis result, we can conclude that our method can basically predict the development of economic and provide specifically economic values for managers.

4.1 Prediction Experimental Result and Comparison

Following Figure 4 demonstrates the trained result and the real situation economic development that measured with Gross National Product (GDP). From following diagram, we can observe that our proposed model can precisely simulate the economic development and provide the reliable predictions.

Additionally, we concentrate the total prediction accuracy that is measured by trained results divided by the real situation. If the accuracy is more than 1 that presenting the current trained value is larger than the real situation and otherwise. Following

From following Figure 5 demonstration, we can observe that our method prediction result representing as orange line is precisely according with the real economic development representing as the blue line that measured with GDP from 2010 to 2019.

The comparison results illustrates that our proposed model performers better accuracy than other existing methods due to other method is much closer to 1 in each prediction iterations.

Subsequently, Figure 5 demonstrates the prediction accuracy that compared with existing two economic prediction method.

4.2 Computation Cost Comparison Results

Another essential indicator is computation cost for the model. Following Table 2 demonstrates the detail computation time cost result compared with mentioned two existing methods.

<table>
<thead>
<tr>
<th>Data Size Level (GB)</th>
<th>Ours</th>
<th>RBF</th>
<th>GM</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>50s</td>
<td>32s</td>
<td>67s</td>
</tr>
<tr>
<td>40</td>
<td>93s</td>
<td>61s</td>
<td>126s</td>
</tr>
<tr>
<td>100</td>
<td>214s</td>
<td>142s</td>
<td>312s</td>
</tr>
</tbody>
</table>
CONCLUSIONS AND FUTURE IMPROVEMENTS

In this paper, we propose a novel structure by utilizing the method of deep convolutional neural network and predict the economic development in the real situation. From our extensive experimental results, we can conclude that our proposed method can be utilized for manager to predict the future economic if the input data is precisely and the number of required data is sufficient. Additionally, from our comparison results, we can observe that our model work with acceptable computation cost and obtain low delay. As for the future improvement, we can dispose the issue about data pro-processing stage and guarantee the model is fairness for each input data and utilize optimization functions to maximize the rewards of neural network.

REFERENCES


Research on Omni-Channel Supply Chain Structure Based on BOPS Mode

Shuang Zhou
shuangasd@hotmail.com
Sultan Idris Education University, Tanjung Malim, Malaysia

Abstract: With the rapid development of Chinese economy, the urbanization process is accelerating and the number of urban population is increasing. In this context, the omnichannel supply chain is a new circulation mode emerged at the historic moment, its development speed is amazing. Under the mode of omni-channel supply chain management, the urban integrated pipeline corridor industry, logistics industry and finance have shown a rapid growth trend and become an important factor to promote the urban economic development. However, there are still many problems in supply chain management in the process of supply-side structural reform. For example, low efficiency of infrastructure construction, high logistics cost and lack of omni-channel service quality lead to low efficiency of supply chain structure, which also affects urban economic development and residents' consumption level.

Keywords: BOPS mode, Omni-channel, Supply chain structure.

1 INTRODUCTION

With the development of the world economy, people's demand for material life and spiritual culture has been increasing, the mode of omni-channel supply chain management has become an indispensable part of enterprise operations, studying BOPS mode under the current environment is of great significance. In the omni-channel mode, studying the relationship between BOPS and the supply chain structure of logistics enterprises can effectively reduce the traditional operation cost and management risk. By analyzing the problems faced by the docking of traditional commercial operation mode and modern logistics industry, it can be found that: On the one hand, disadvantages such as large investment in infrastructure construction and long cycle limit the transformation of traditional domestic retail industry to modernization\cite{1}. On the other hand, from the perspective of the government, it proposes to introduce competition mechanism into the supply chain, improve the level of omni-channel management and service efficiency, so as to promote the strategic theoretical basis of enterprise development. At the same time, this will help solve the problems in the process of docking Chinese traditional mode of operation and modern logistics industry and improve the competitiveness of enterprises.
2 RESEARCH ON SUPPLY CHAIN DECISION-MAKING UNDER DIFFERENT CHANNEL STRATEGIES

2.1 Problem Description

Omnichannel means that in the process of supply chain management, enterprises, consumers and retailers cooperate to sell their own products to end users. Theoretically speaking, it is a management mode in which suppliers, manufacturers and distributors jointly complete value-added and after-sales service activities to achieve social and economic benefits. This is the concrete application of the concept of "omnichannel" in the field of logistics, but in practice, there are still problems such as unreasonable supply chain structure and low operating efficiency. At present, the main reasons for the slow development of Chinese traditional retail industry are: the lack of core technical strength; On the other hand, due to the lack of large-scale production, it is unable to form large-scale operation advantages and long-term profitability, which leads to a disadvantageous position in market competition, this is difficult to achieve a sustainable virtuous cycle.

Consider the secondary supply chain composed of traditional retailers and manufacturers, and assume that there are two modes: dual channel and omnichannel: (1) Dual channel refers to the supply chain formed by retailers and manufacturers through franchising, that is, retailers establish contact with wholesale and retail enterprises through the construction of wholesalers, channel agents, sales agents and other modes[2]. (2) Omnichannel means not only retailers' physical stores and manufacturers' online direct selling channels, but also the online direct selling channels established with e-commerce, that is, retailers establish contact with wholesale and retail enterprises by establishing physical stores and sales agents of wholesale and retail enterprises. Retailers build sales agents and other models to form supply chains with them. As can be seen from the above description, the channel structure under dual-channel and all-channel (i.e. channel under BOPS mode) is shown in Figure 1:

![Figure 1. Dual-channel and all-channel channel structure](image)

In a dual-channel supply chain structure, the retailer first purchases the product from the manufacturer at the wholesale price \( w \), and then the retailer ships the product to the retailer. Then the retailer decides the offline retail price of the product \( p_s \) to sell the product and provides the service value \( e \) for consumers in the physical store[3]. It also provides logistics support for the final sales process and ensures the integrity of the product supply chain. Therefore, how to coordinate the relationship between retailers and wholesalers, distributors and consumers is a
topic that must be faced to realize the value maximization of dual channels. In this mode, manufacturers, distributors and third-party logistics enterprises constitute an important part of the whole omnichannel system structure. The optimization of the supply chain structure is the key to maximize the omnichannel value and improve the overall efficiency and economic benefits of the whole system. Market demand depends on the selling price and the service value of consumers. Therefore, under the dual-channel supply chain structure, the demands of traditional channels and online direct selling channels are respectively shown in formula (1) and (2):

\[ D_s = (1 - \lambda) a - (p_s - e) + \theta p_o \]  

\[ D_o = \lambda a - p_o + \theta (p_s - e) \]  

In addition to the above dual-channel supply chain structure, in order to comply with the integration trend of online and offline channels, assuming that manufacturers and retailers cooperate to provide BOPS channels, this paper studies the omnichannel composed of manufacturers and retailers cooperate to provide BOOTC products under the traditional supply chain[4]. Based on the analysis, the optimization scheme of the all-channel supply structure based on the BOPS mode is obtained. By analyzing the comparison between the traditional supply chain and the online and offline channels, this paper concludes that the optimization of the all-channel supply structure plays an important role in improving the overall operating efficiency and reducing the logistics cost. Research shows that 45% of consumers who use BOPS channels purchase additional products during in-store pickup, while those purchased in retail stores are composed of online and offline channels; Consumers can collect goods from merchants by means of online payment. Secondly, retailers differ from traditional supply chains because of information sharing and service demand differences between online and stores, which make it difficult to effectively connect them[5]. Therefore, more and more retailers are willing to adopt the BOPS strategy. In the case of considering BOPS channels, the demand function model of the omnichannel structure is constructed as shown in formula (3), (4) and (5):

\[ D_s^b = (1 - \lambda) a - (p_s - e) + \theta p_o \]  

\[ D_o^b = m(\lambda + \delta) a - p_o + \theta (p_s - e) \]  

\[ D^c = (1 - m)(\lambda + \delta) a - (p_s - e) + \theta (p_s - e) \]

2.2 Optimal Decision Making Under Centralized Supply Chain

The optimal decision under the centralized supply chain means that the supplier bears the cost of the overall structure of the supply chain and shares the risk through third-party logistics enterprises or financial institutions according to customers’ new requirements on the performance, quality and service of the goods under the condition of ensuring timely supply and
demand. It can reduce the overall demand level of both parties and improve customer satisfaction and service quality[6]. Under the centralized supply chain structure, the traditional retail channel and the online channel are managed by a central decision maker. The analysis of the dual channel and omnichannel models is as follows:

Model analysis without BOPS channels: Based on the description of dual-channel demand, related concepts of supply chain management and omnichannel are analyzed. From the perspective of demand side, combined with the BOPS model of the coordination mechanism between channels, operational efficiency and other influencing factors, the conclusion is drawn: with the intensification of market competition and the reductio n of industry access threshold, the profit space of enterprises will be compressed; On the other hand, the change of consumer consumption habits leads to the increase of retail terminal diversion. In the case of centralized decision-making, the profit of the central decision maker operating the dual-channel supply chain is expressed as Formula (6):

$$\pi_c = p_s[(1-\lambda)a - (p_s - e) + \theta p_o] + p_o[\lambda a - p_o + \theta (p_s - e)] - \frac{\beta e^2}{2}$$  \hspace{1cm} (6)

In the case of centralized decision making, the profit of the central decision maker consists of two parts: one is the profit of selling products in traditional physical stores, and the other is the profit of franchise sales. Central decision makers play an important role in supply chain management, whose main function is to analyze and optimize the whole system structure, which can effectively control and supervise the overall operation mode[7].

Model analysis with BOPS channels: In the practice of omnichannel retail, BOPS strategy greatly improves the convenience of consumers to buy products and the flexibility of order fulfillment, which improves the overall efficiency of the supply chain and provides consumers with more high-quality and convenient shopping experience, thus improving the omnichannel service quality. In the case of centralized decision-making, the study of omnichannel supply chain structure is conducive to improving the enterprise's own resource allocation and market competitiveness, and promoting industrial upgrading. Combined with the demand function of the omnichannel supply chain, the profit function of the central decision maker under the omnichannel supply chain can be obtained as shown in Formula (7):

$$\pi_c^o = p_s[(1-\lambda)a - (p_s - e) + \theta p_o] + p_o[(\lambda + \delta) a + 2(\theta p_s - p_o - \theta e) + e] - \frac{\beta e^2}{2}$$  \hspace{1cm} (7)

2.3 Balanced Decision Making in Decentralized Supply Chain

In the decentralized mode, the retailer's decision-making power is mainly in the hands of the supplier, so it needs to be supervised and managed. When a retailer makes a purchase decision, the supply chain system will formulate corresponding strategies according to different demands, which leads to adverse selection problems due to consumer dispersion, information asymmetry and other reasons. On the other hand, it may also lead to the failure of the manufacturer to obtain sufficient product supply, resulting in increased costs and even a loss operation situation.
Meanwhile, the scattered distribution of consumers makes it difficult for the enterprise to centrally control the type and quantity of goods. Therefore, supervision and management should be carried out to achieve the goal of reducing risks and improving profits\[8\].

Model analysis of non-BOSP channel: The model of BOSP channel is established under the traditional distribution model, and enterprises need to settle the related expenses involved in the process of product production to sales. This enables decision-makers to have a certain understanding of their regional market environment. Therefore, in order to solve this problem, we must first determine a reasonable and feasible scheme; Secondly, it is necessary to formulate corresponding strategies and measures according to consumer demand preferences in different regions. Finally, considering the cost factor and combining with the actual situation to make the most suitable distribution channel selection, so as to reduce the cost of decision makers and improve the profit margin. Therefore, under the set and distributed supply chain, the profit of the retailer operating the dual-channel supply chain is expressed as formula (8):

\[
\pi_r = (p_r - w)((1 - \lambda)\alpha - (p_c - e) + \theta p \lambda) - \frac{\beta e^2}{2}
\]

Under the centralized distributed supply chain, the profit of the central decision maker consists of two parts: one is the profit of selling products in traditional physical stores, and the other is the profit of franchise sales \[9\]. Central decision maker plays an important role in supply chain management. Its main function is to analyze and optimize the whole system structure, and it can effectively control and supervise the overall operation mode.

Model analysis when there are BOPS channels: Nowadays, consumers' purchasing methods are gradually changing to multi-channel, mobile and fragmented, and the consumption process is becoming simpler and more transparent. Therefore, the development of enterprises faces great challenges. In such a complex competitive environment, it is necessary to make appropriate adjustments in order to remain invincible, and the equilibrium strategy means to achieve the overall optimal through the interaction between different types and scale factors. Decision makers need to constantly revise strategic selection methods according to market changes, and make timely responses to meet consumer demand and improve customer satisfaction and loyalty. Therefore, the profit function of omnichannel supply chain in decentralized decision-making is expressed as Formula (9):

\[
\pi^*_c = (1 + \delta)a - (1 + \delta)m + \theta(p_r + p_c) + (2 - \theta)e
\]

3 OMNICHANNEL SUPPLY CHAIN STRUCTURE BASED ON BOPS MODE

3.1 Construction of Omnichannel Supply Chain Structure

Omnichannel supply chain structure refers to the vertical connection between suppliers and final consumers, and its core is to build a system composed of manufacturers, distributors and
retailers. Through the establishment of the supply chain management organization to ensure that the enterprise in the whole operation process to the market changes in time, and through the supply chain management to achieve rapid response to the market demand, so as to ensure the quality of products and service level. The omnichannel supply chain model in BOPS mode is shown in Figure 2:

![Figure 2. Omnichannel supply chain model under BOPS mode](image)

In the practice of omnichannel retail, the integration of online and offline channels not only needs to overcome many technical bottlenecks, but also needs good cooperation and collaboration among supply chain members. Under the BOPS mode, the effective integration of online and offline channels is conducive to making up for the defects of the traditional supply chain management mode, such as low efficiency, high cost and long cycle. On the one hand, BOPS increases the sales and service cost of offline retail stores and also increases the sales profit of online channels. On the other hand, for offline retail stores, cooperation with online platforms can enable them to obtain more professional and systematic services. At the same time, in terms of BOPS sales integration, enterprises have adopted two strategies in practice, including the offline channel and the online channel. However, due to the high supply chain management cost of the online platform and the limitation of the offline channel, enterprises are faced with great difficulties in the operation process.

### 3.2 BOPS Sales are Credited to Online Channels

Omni-channel management based on channel structure refers to the income and expenditure generated from a supply chain and gradually refined to other links and processes in the whole value chain. The inclusion of BOPS sales into online channels means that all the profits of BOPS consumer orders belong to manufacturers, and the benefits of consumers through channels are obtained by manufacturers. Supply chain structure will directly determine the value creation ability and efficiency. At this time, the manufacturer's profit optimization is expressed as Formula (10):

$$\max_w \pi_m^{D1} = wD_s + p(D_b + D_d) = w(\lambda a + e - p) + p[(1 - \lambda) a + e + 2p]$$ (10)
When BOPS sales volume is included in online channels, it brings huge economic benefits to the current society. At the same time, it also increases the government's investment in infrastructure construction and improves the overall competitiveness of the city\(^{(10)}\).

### 3.3 BOPS Sales are Booked into Offline Channels

After studying the omnichannel supply chain structure, it is found that it is mainly operated through online sales mode. Offline channel refers to that retailers directly purchase products from manufacturers and charge consumers fees when purchasing goods or services from wholesalers. However, there are huge differences between traditional retail enterprises and e-commerce: on the one hand, they have different places of operation, on the other hand, there are obvious differences - the limitations of the point market and the terminal market, so the online sales model to a large extent limits the cost and risk sharing of business exchanges between the main operators of the channel. In the omnichannel practice, in order to improve the service enthusiasm of offline retailers, improve customer satisfaction, and promote the healthy development of online retail enterprises, this paper studies the omnichannel supply chain when BOPS sales volume is included in the offline channel, and the profit optimization problem of the manufacturer is shown in Formula (11):

$$\max_w \pi^o_w = w(D_s + D_o) + pD_o = w[(\lambda_1 + \lambda_2)a + 2e - 2p] + p[(1 - \lambda_1 - \lambda_2)a - p](11)$$

When BOPS sales are included in offline channels, its advantage is that it can improve the efficiency of the supply chain, reduce operating costs, and increase the channel supply capacity to a certain extent.

### 4 CONCLUSION

In summary, with the development of social economy and improvement of people's living standards in our country, the concept of omni-channel supply chain is widely applied in all walks of life, and the theory study and practice exploration of supply chain management model also gradually increase. Our country carries on an active response under the field of omni-channel, which is of great significance for promoting the development of the society. However, at present, domestic academic circles have few researches on the omnichannel logistics structure under the BOPS mode, and most of them focus on its management and control in traditional urban construction, while the research on omnichannel supply chain management of urban infrastructure is still in its infancy. Based on this, this paper studies the omnichannel supply chain structure under the BOPS mode, hoping to provide some references for related fields.

### REFERENCES


Corporate Performance Prediction Based on BP Neural Network

Hongyan Ye
yehongyan@stu.scu.edu.cn

Business School, Sichuan University, No.24 South Section 1, Yihuan Road, Chengdu, China

Abstract: Top Management Team (TMT), as the most important presence in corporate decision-making, is an integral part of corporate research. Its impact on corporate performance is a topic that cannot be ignored. This paper uses earnings per share (EPS) to represent corporate performance, and the shareholding ratio of executives (SHARE) is used as TMT characteristics. Based on the BP neural network, the input layer of the model is set as five nodes, the implied layer as two nodes, and the output layer as one node. We use 75% of the data of listed information technology companies from 2017-2019 as the training set to derive the performance prediction model. In this study, 25% of the test set is used to validate the final valid performance prediction model obtained. This study integrates TMT characteristics into predicting corporate performance, helping to optimize the non-economic indicators used to assess and predict corporate performance.

Keyword: Top Management Team Characteristics, R&D Investment, Corporate Performance, BP Neural Network.

1 INTRODUCTION

The top management team (TMT), as the makers and executors of corporate decisions [2], has been a high research topic around the world (Hambrick and Mason, 1984). Since the upper echelons theory was proposed, many studies have been conducted by domestic and international scholars on the demographic characteristics of TMT and Pay Dispersion in TMT on corporate performance [3][6]. However, the results of TMT characteristics and corporate performance have not reached unity, and numerous theories have been born to explain their relationship [1].

With the advent of the 5G era and the IoT technology revolution, people are paying more attention to R&D, and R&D investment is increasingly reflecting a country’s level of economic development. According to the “2020 National Science and Technology Expenditure Statistics Bulletin” published by the National Bureau of Statistics, the Ministry of Science and Technology, and the Ministry of Finance in 2021, China’s R&D investment has increased in 2020 and has maintained a double-digit growth rate since the “13th Five-Year Plan”, a record high growth rate.

Previous studies have tended to focus on economic indicators and ignore the impact of TMT [8]. Therefore, this paper focuses on the prediction model of earnings per share (EPS) by adding TMT characteristics variables and using information technology-listed companies as research samples.
2 THEORETICAL BACKGROUND AND HYPOTHESES

TMT, as the most important presence in corporate decision-making [2], is an integral part of accounting research and its impact on corporate performance is a topic that cannot be ignored. According to the most popular explanation of agency theory, in the case of information asymmetry, it is the executive team members as *Economic Men*. Individuals will constantly seek to maximize their own interests, making them inconsistent with the interests of the firm, leading to free-riding [9][11].

Agent theory suggests that increasing the shareholding of TMT will reduce this goal inconsistency (Garvey, 1992). Giving shares to the executive team is a form of equity incentive. The higher the shareholding ratio, the more it reduces the short-sighted effect of the executive team on R&D investment activities and motivates the executive team to invest in R&D. In the information technology industry, a large part of a company’s performance relies on its degree of technological innovation. It is not difficult to imagine that the higher the intensity of R&D investment, the stronger the technological innovation. In summary, we posit that Executive team shareholding, R&D investment increase corporate performance.

3 METHOD AND MATERIAL

3.1 BP Neural Network

BP neural network is a multiple-feedforward network trained according to the error backpropagation algorithm [7]. It can store and learn a large amount of input and output data by simulating the function of human neurons. It does not need to describe the mapping relationship of variables, using input and output data to the model. It has a solid ability to simulate nonlinear systems. The BP neural network often consists of input, output, and hidden layers. The neurons between layers are connected in a fully interconnected manner, interconnected by corresponding network weight coefficients. The neurons within each layer are not connected. See Figure 1 for an illustration.

![Figure 1: BP neural network topology.](image-url)
The BP neural network training process is as follows. We suppose that there are \( A \) training samples to train the network, one of which is \( a \). For \( a \), the input to the neuron \( n \) in layer \( i \) is as follows.

\[
\text{net}_{na}^{(i)} = \begin{cases} 
  x_n, & i = 1 \\
  \sum_{j=1}^{N_{i-1}} w_{nj}^{(i)} o_{ja}^{(i-1)} - o_{n}^{(i)}, & i = 2
\end{cases}
\]  

(1)

Where \( x_n \) is the input of the neuron \( n \), \( w_{nj}^{(i)} \) is the connection of weight between neuron \( n \) of the layer \( i \) and neuron \( j \) of the layer \((i-1)\) in equation (1). \( o_{ja}^{(i-1)} \) is the output of neuron \( j \) of the layer \((i-1)\) of sample \( a \) in equation (1).

The output of the neuron \( n \) of the layer \( i \) is as follows.

\[
\hat{o}_{na}^{(i)} = \begin{cases} 
  \text{net}_{na}^{(i)}, & i = 1 \\
  f(\text{net}_{na}^{(i)}), & i = 2
\end{cases}
\]  

(2)

Where \( f(.) \) is the activation function in equation (2).

### 3.2 Sample Selection

In this paper, the listed companies in the information technology industry in the SSE A-share and SZSE A-share in 2017 year and 2018 year are used as the research sample. The data on corporate performance is obtained from the 2018 year and 2019 year data because of the lagging effect of the growth effect of corporate performance brought about by considering R&D investment.

According to the collected samples, the following treatments were performed: (1) eliminating samples with missing relevant variables; (2) eliminating samples of ST and *ST enterprises; (3) Winsorize all variables from 1% to 99% in order to remove the influence of extreme values, and finally, 1045 samples were obtained. The data were all obtained from the CSMAR database.

### 3.3 Input Layer, Hidden Layer, Output Layer

In this study, 75% of the data were used as training samples and 25% as test samples. The input layers are SHARE, RD, Size, Growth, and AGE and the output layer is EPS.

- **SHARE**: The shareholding ratio of executives is represented by the ratio of the number of shares held by executives to the total number of shares of the company.
- **RD**: R&D investment can be measured by the intensity of R&D investment.
- **Size**: Corporate performance (EPS): We collect EPS for 2018-2019 as the dependent variable for this study. Company size (Size): Liu and Liu proposed that company size is closely related to research investment.
- **Growth**: Companies with higher growth will care more about their corporate performance the following year.
- **AGE**: The older the company, the greater stability of executive team characteristics such as executive shareholding ratio. Table 1 shows the variable descriptions.
There is no definite formula for calculating the number of stages of the hidden layer. Therefore, when selecting an implicit layer, its reasonable range of values in equation (3) is usually calculated based on an empirical formula.

\[ k = \sqrt{m + n} - a, \ (0 < a < 10) \]  

(3)

Where \( k \) is the number of nodes in the hidden layer, \( n \) is the number of nodes in the input layer, and \( m \) is the number of nodes in the output layer. Therefore, the number of nodes in the hidden layer in this study is 2.

### 3.4 Model Design

We construct a multiple linear regression model and use BP neural network to test. We further establish equation (4).

\[ EPS = a_0 + a_1 SHARE + a_2 RND + a_3 Size + a_4 Growth + a_5 AGE + \varepsilon \]  

(4)

### 4 RESULT

#### 4.1 Data Processing

In this study, we cleaned the data and tested and plotted the hypotheses using Stata 17.0 and Python 3.8.

Table 2 presents the results of descriptive statistics for all variables. Figure 2 presents that the data which correspond to normal distribution were analyzed.

<table>
<thead>
<tr>
<th>Variables</th>
<th>N</th>
<th>Mean</th>
<th>S. E</th>
<th>Mini</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>SHARE</td>
<td>1045</td>
<td>0.119</td>
<td>0.158</td>
<td>0</td>
<td>0.662</td>
</tr>
<tr>
<td>RD</td>
<td>1045</td>
<td>0.044</td>
<td>0.032</td>
<td>0.000</td>
<td>0.160</td>
</tr>
<tr>
<td>EPS</td>
<td>1045</td>
<td>0.312</td>
<td>0.986</td>
<td>-3.280</td>
<td>3.68</td>
</tr>
</tbody>
</table>
Each variable has a different physical unit and represents a different economic significance. In this paper, to avoid large differences between the values of the variables, we have dimensionless processed the sample data according to the equation (5) and equation (6) as follows.

\[
x_{ij} = \frac{(\text{var}_{ij} - \text{min}(\text{var}_{in}))}{(\text{max}(\text{var}_{in}) - \text{min}(\text{var}_{in}))}
\]

(5)

\[
x_{ij} = \frac{(\text{max}(\text{var}_{in}) - \text{var}_{ij})}{(\text{max}(\text{var}_{in}) - \text{min}(\text{var}_{in}))}
\]

(6)

The results of the treatment are shown in Table 3. Furthermore, the results of correlation analysis are shown in Table 4.

**Table 3: Dimensionless processing data.**

<table>
<thead>
<tr>
<th>Variables</th>
<th>N</th>
<th>Mean</th>
<th>S. E</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>SHARE</td>
<td>1045</td>
<td>0.188</td>
<td>0.239</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>RD</td>
<td>1045</td>
<td>0.258</td>
<td>0.179</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>EPS</td>
<td>1045</td>
<td>0.412</td>
<td>0.067</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Size</td>
<td>1045</td>
<td>0.364</td>
<td>0.151</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Growth</td>
<td>1045</td>
<td>0.019</td>
<td>0.049</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>AGE</td>
<td>1045</td>
<td>0.276</td>
<td>0.143</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 4: Results of correlation analysis.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>EPS</td>
<td>1</td>
<td>0.235***</td>
<td>1</td>
</tr>
<tr>
<td>RD</td>
<td>0.157***</td>
<td>0.126***</td>
<td>-0.245***</td>
</tr>
<tr>
<td>SHARE</td>
<td>-0.021</td>
<td>-0.072*</td>
<td>-0.031</td>
</tr>
<tr>
<td>Size</td>
<td>-0.010</td>
<td>-0.082*</td>
<td>-0.180***</td>
</tr>
<tr>
<td>Growth</td>
<td>0.101**</td>
<td>1</td>
<td>-0.054</td>
</tr>
<tr>
<td>AGE</td>
<td>-0.0280</td>
<td>-0.082*</td>
<td>-0.180***</td>
</tr>
</tbody>
</table>

Furthermore, we use the Stata17.0 graphing tool and Python 3.8 to present the correlation between SHARE, R&D, EPS more clearly in the figures. Figure 3 shows the relationship between EPS and SHARE, which is the plotted result of Model 1, and the relationship between R&D and SHARE, which is the plotted result of Model 3. A 3D result plot of SHARE, R&D, and EPS is shown, demonstrating a visible surface in Figure 4, which is the plotted result of SHARE-R&D-EPS relationship.

![Figure 3: The correlation between SHARE and EPS and the correlation between SHARE and RND.](image-url)
4.2 Training and Testing of Neural Network Predictive Model

In this paper, the model is analyzed using Stata 17.0, and the input layer is set to 5 nodes, the hidden layer is set to 2 nodes, and the output layer is set to 1 node. BP neural network is established after 200 iterations, and the error of its model reaches the requirements of the set criteria, that is, the corporate performance-based prediction model. The error of the network training is shown in Figure 5.

The “Tansig” function is selected for the activation function of the neurons in the hidden layer of the first layer and the neurons in the output layer of the second layer.

In addition, we compared the R-squared of the BP neural network-based model with that of the OLS-based model. It is found that the BP neural network model R-squared is significantly larger than the OLS method model, indicating that the BP neural network method prediction method is better than the OLS method. See Table 5 for more details.

<table>
<thead>
<tr>
<th>Method</th>
<th>R-Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>BP Neutral Network</td>
<td>0.78</td>
</tr>
<tr>
<td>OLS</td>
<td>0.25</td>
</tr>
</tbody>
</table>
In order to test whether the BP neural network predictive model has good generalization ability, the model also needs to be tested. Using the remaining 25% of the data, we performed the test. The test results are shown in Figure 6.

**Figure 6:** Test results of BP neural network predictive model.

From the Figure 6, the model trained using 75% of the data in this study as training can predict the remaining 25% of the corporate performance level very well. The relative errors of the test sets are all relatively small, implying that this BP neural network model has good generalization ability and can have good corporate performance prediction for Chinese IT-listed companies.

5 CONCLUSION AND FUTURE RESEARCH DIRECTION

5.1 Conclusion

In this paper, we add top management team characteristics to predict corporate performance, which only remedies the previous evaluation of corporate performance around economic indicators. In addition, neural networks are utilized to optimize the construction of multiple linear models compared to OLS. Furthermore, this study integrates TMT characteristics into predicting corporate performance, helping to optimize the non-economic indicators used to assess and predict corporate performance.

This study establishes a BP neural network predictive model for testing the prediction of EPS. The findings verify the positive impact of the shareholding ratio of the top management team on corporate performance. This study reveals that the amount of the principal’s shareholding in the executive team as an equity incentive aligns the agent with its goals to make long-term, correct investments in R&D investment, increasing corporate performance.

From the perspective of Chinese culture, compensation represents having a higher income and a status symbol. The shareholding ratio can also bring higher satisfaction to the higher-level executive members, and the decisions are more in the hands of the executives at the higher compensation levels, so increasing the shareholding ratio of the executive team will increase corporate performance.
5.2 Future Research Direction

This study has considered the effect of top management team characteristics, but ignored the full range of economic indicators. Future research should focus on integrating top management team characteristics and economic indicators to explore the prediction of corporate performance evaluation.

In addition, this study only regarded the data before the epidemic. Future research could adopt the DID (Differences-in-Differences) method and consider whether there is a change in the prediction for corporate performance evaluation in the epidemic context.

REFERENCES

A Hybrid Model Integrating LSTM with Multiple GARCH-Type Models for Volatility and Var Forecast

Jiayi Liu*

*Corresponding author’s e-mail: jiayiliu27@gmail.com

Financial Mathematics, NC State University, Raleigh, NC, 27607, the United States

Abstract: The prediction of volatility and Value-at-Risk is a key problem in finance which can help to measure the risk, or the error sizes obtained in modelling several financial variables. In this study, I use high-frequency data to calculate the realized volatility which will be implemented as a financial asset volatility measurement. Furthermore, a hybrid model that integrates three GARCH-type models and LSTM (Long Short-Term Memory) neural networks is proposed to forecast the volatility of CSI 300 index and further translate to accurate VaR (Value at Risk). Then the performance of the hybrid model prediction is compared against the performance of standalone models such as LSTM with RV as the only input. In the end, I generate the trading strategy signal and built the strategy for investment to explore the applicability of this hybrid model in risk management. The empirical analysis of the above model is carried out on CSI 300 index data. The empirical results in this study showed that the hybrid model could significantly improve the volatility and VaR prediction performance of the CSI 300 Index. Therefore, the research methods and the conclusions of this study could provide the possibility for the further application of this hybrid model in financial research.

Keywords: Volatility, GARCH-type models, LSTM model.

1 INTRODUCTION

COVID-19 has been spread all over the world since 2020. The global economy has suffered badly as a result of the pandemic. The CSI 300 Index which is thought to be the ‘Blue Chip’ index for the mainland China stock exchange got a maximal drawdown of 33.52% since 2021, the largest maximal drawdown since 2015. The S&P 500 Index plunged five times between February and March 2020, leading to a market meltdown. Many investors also suffered losses in the abnormal volatility of the financial market, so the tail risk of asset return under extreme volatility has become the focus of research. Volatility plays an important role in many fields of finance, for example, derivative pricing, portfolio risk management, hedging strategies and systemic risk. Therefore, it is worthwhile for investors to make better use of volatility information to construct trading strategy.

The significance of this paper is as follows: (1) Using multiple models studying RV to improve accuracy and robustness of volatility predictions; (2) Doing out-of-sample forecast to evaluate the model performance; (3) Using VaR estimates to do risk analysis; (4) Combining artificial intelligence algorithm and traditional volatility model to improve model performance; (5)
Providing a reference model for investment and risk management which could promote market pricing efficiency and stability.

The remainder of this paper is organized as follows. In Section 2, I review the previous related literature. Section 3 outlines the basic models and methods used in this paper. In Section 4, I introduce data for experiment and show the process of experiment. Section 5 outlines the results of the experiment. And the conclusion is given in Section 6.

2 LITERATURE REVIEW

Engle (1982) introduced autoregressive conditional heteroscedastic (ARCH) to estimate the means and variances of inflation in the U.K. Based on Engle’s study, Andersen and Bollerslev (1986) put up generalized Autoregressive Conditional Heteroskedastic (GARCH) which provided a more rational lag structure. Then Nelson (1991) put forward EGARCH model which addresses conditional heteroscedasticity, or volatility clustering, in an innovations process. The GARCH-type models could capture the characteristic of the in-sample volatility well in financial time series data. Most of previous work in this area using these linear models. These models mentioned above are traditional time series models which are easy to interpret in statistics and has been widely used for estimating the volatility of the financial sector for years. But they all contain stationary assumptions which is hard to satisfy by the financial data in the real world, which makes it hard to get a desirable result on out-of-sample data using GARCH-type models’ empirical properties. To improve this, Andersen and Bollerslev (1998) first put up the concept of realized volatility (RV), which does not depend on the specific assumptions taken by model used to measure the volatility and reduce the measurement errors by using high frequency data. Shao and Yin (2008) built up a realized volatility model and a realized range model to finally compute VaR (Value at Risk) by using intraday high-frequency data. This gives an example that prove models based on intraday data are better performed than models based on daily returns. RV could display the volatility that could not be observed before and can measure the fluctuation of high-frequency data. It can be confirmed that the accuracy of the estimate improved by using high-frequency data. Consider this, it is reasonable to think of RV calculated by high frequency data as the actual volatility in this study.

With the development of artificial intelligence and increased computational capabilities, people began to implement powerful machine learning method in financial time series modelling, such as stock price prediction. Method such as support vector machine (SVM), Random Forest (RF). Some methods are based on neural networks such as artificial neural network (ANN), Convolutional Neural Network (CNN), Recurrent Neural Network (RNN) and deep neural networks like Long Short-Term Memory (LSTM). Barunik and Krehlik (2016) first used ANN to predict the volatility in the energy market, and improved accuracy of prediction by using high frequency data. ANN is a type of machine-learning algorithm and a data-driven nonparametric method. As time goes by, a huge amount of financial data become accessible, which makes ANN an ideal method when enough data exists. Hochreiter and Schmidhuber (1997) first introduced Long Short-term Memory (LSTM) algorithms which is the most famous form of Recurrent Neural Network to solve complex, long time lag tasks efficiently. Unlike traditional predictive learning models, RNN collects memory of the data path and expose features of hidden states to track statistical patterns. This avoids calibration of numerous macroeconomic and
company-specific variables in forecasting stock prices. The end-to-end mapping system brings the convenience of nonparametric statistical inference. In practice, people usually choose LSTM to avoid long-term dependency problems. Chen, Zhou, and Dai (2015) used Long Short-Term Memory (LSTM) to predict China stock returns, which proves the possible use of LSTM in stock market prediction. In this study, I will explore the applicability of Long Short-Term Memory (LSTM) to predict volatility.

Recent work shows that stock market prediction could be improved using a hybrid model. Kim and Won (2018) put up a hybrid model combing LSTM with multiple GARCH-type models to forecast the realized volatility of the KOSPI 200 index and showed that the hybrid model performs better than any other single GARCH-type model. Kuster et al. (2006) showed the accuracy prediction of volatility is of great importance to predict VaR. In this study, I will try to use different hybrid models to forecast volatility and extend their study in risk analysis.

3 METHODOLOGY

3.1 Realized Volatility

In this study, my goal is to compare predicted versus actual volatility, which is set as the target value for the supervised learning process. The following method I used to calculate Realized Volatility (RV) is based on the research by Andersen and Bollerslev (1998). To calculate the daily realized volatility of the \( t \)-th day,

\[
RV_t^d = \sum_{i=1}^{M} r_{t,i}^2
\]

\( r_{t,i} = 100 \times \ln \left( \frac{p_{t,i}}{p_{t,i-1}} \right) \), denotes the i-th close price in the t-th day.

\( p_{t,i} \) denotes the i-th day close price

\( M \) denotes sample frequency.

This realized volatility is used as the actual volatility.

3.2 Model

3.2.1 Standard GARCH Model

Bollerslev (1986) put forward GARCH model, this model is equivalent to ARCH-infinite model. The standard GARCH(1,1) model is:

\[
y_t = \varphi x_t + \mu_t, \mu_t \sim N(0, \sigma_t^2)
\]

\[
\sigma_t^2 = Var(y_t|I_{t-1}) = \alpha_0 + \alpha_1 \mu_{t-1}^2
\]
where $y_t$ is a given stochastic term, and $\mu_t$ is the drift. $\sigma_t^2$ denote the volatility at time $t$. $\sigma_t^2$ means that given information before time $t$. $N(0, \sigma_t^2)$ denote the standard Gaussian distribution. All coefficients in the equations above are set to be non-negative.

### 3.2.2 Exponential GARCH Model

Exponential GARCH (or EGARCH) model was put up by Nelson (1991) to overcome some weakness of GARCH model in financial time series. Compared to GARCH model, the coefficients in the EGARCH model could be negative and this model can show the leverage effect, which reflects the asymmetric impacts of negative and positive impacts of the same magnitude. The form for EGARCH(m,s) model is:

$$\ln(\sigma_t^2) = \alpha_0 + \sum_{i=1}^{s} \frac{|\alpha_{t-i} + \gamma_i \epsilon_{t-i}|}{\sigma_{t-i}} + \sum_{j=1}^{m} \beta_j \ln(\sigma_{t-j}^2)$$ (4)

Here a positive $\alpha_i$ contributes $\alpha_i(1 + \gamma_i)|\epsilon_{t-i}|$ to the log volatility, whereas a negative $\alpha_{t-i}$ gives $\alpha_i(1 + \gamma_i)|\epsilon_{t-i}|$, where $\epsilon_{t-i} = \frac{\alpha_{t-i}}{\sigma_{t-i}}$. The $\gamma_i$ parameter thus signifies the leverage effect of $\alpha_{t-i}$. Again, we expect $\gamma_i$ to be negative in real applications.

### 3.2.3 Threshold GARCH Model

The threshold GARCH (or TGARCH) Model proposed by Runkle (1993) and Zakoian (1994) was designed to handle leverage effects. A TGARCH(m,s) Model assumes the form:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{s} (\alpha_i + \gamma_i N_{t-i}) \alpha_{t-i}^2 + \sum_{j=1}^{m} \beta_j \sigma_{t-j}^2$$ (5)

where $N_{t-i}$ is an indicator for negative $\alpha_{t-i}$, that is,

$$N_{t-i} = \begin{cases} 1, & \alpha_{t-i} < 0 \\ 0, & \alpha_{t-i} \geq 0 \end{cases}$$ (6)

and $\alpha_i$, $\gamma_i$, and $\beta_j$ are nonnegative parameters satisfying conditions similar to those of GARCH models. From the model, it is seen that a positive $\alpha_{t-i}$ contributes $\alpha_i \alpha_{t-i}^2$ to $\sigma_t^2$, whereas a negative $\alpha_{t-i}$ has a larger impact $(\alpha_i + \gamma_i)\alpha_{t-i}^2$, with $\gamma_i > 0$. The model uses zero as its threshold to separate the impacts of past shocks.

### 3.2.4 Long Short-Term Memory

RNN (Recurrent Neural Network) is used to predict sequential data. RNN consists of input, hidden, and output layers. Classical RNN has a disadvantage which is the vanishing gradient problem. LSTM is designed to deal with this problem.

The feed-forwarding process of LSTM for the input data $x_t$ and hidden state $h_t$ at time-step $t$ can be formulated as follows:
\[ i_t = \sigma(W_iX + b_i) \]  
\[ f_t = \sigma(W_fX + b_f) \]  
\[ o_t = \sigma(W_oX + b_o) \]  
\[ g_t = \tanh(W_gX + b_g) \]  
\[ c_t = c_{t-1} \times f_t + g_t \times i_t \]  
\[ h_t = \tanh(c_t) \times o_t \]

Where $W_i$ and $b_i$ are weights and bias terms, respectively, and $X = \begin{pmatrix} h_t \\ h_{t-1} \end{pmatrix}$. Function $\sigma$ and $\tanh$ are defined by $\sigma = \frac{1}{1 + e^{-x}}$ and $\tanh = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}$.

### 3.2.5 Proposed Hybrid Model

Many studies have demonstrated that the combination of neural networks and GARCH-type models could improve prediction accuracy of volatility compared to using GARCH-type models only. Kim and Won (2018) showed that using information extracted by multiple GARCH-type models as inputs could get a better performance than by one GARCH model. The GARCH model could capture volatility clustering and leptokurtosis information, EGARCH model is used for leverage effect modeling. Hence, each GARCH-type model has different focuses and functions in its volatility prediction. For this reason, combining two or more GARCH-type models could reflect different time series characteristics, which will bring more information. Since Wiśniewska and Wyłomańska (2017) showed that Generalized Error distribution is more adequate to real financial time series than classic Gaussian distribution, a Generalized Error distribution is used for the GARCH models. The final hybrid model consists of 4 models: RV-LSTM, sGARCH-LSTM, eGARCH-LSTM, tGARCH-LSTM, and take the average value as the final volatility estimates.

### 4 EXPERIMENT

#### 4.1 Data

The historical trading data of CSI 300 Index used for the experiment in this study is obtained from JoinQuant. CSI 300 Index is designed to replicate the performance of the top 300 stocks traded in Shanghai Stock Exchange and Shenzhen Stock Exchange. As shown in Figure 1, this dataset consists of 68,976 5-minute data points and 1,436 daily data points from August 23, 2016 to July 22, 2022. When training the LSTM model, 90% of the data in the training set was used as the holdout set to fit the model, and 10% was used as the validation set to tune the hyperparameters.
The following table 1 shows the descriptive statistics of the return and RV (Realized Volatility) of CSI 300 Index such as mean, standard deviation, skewness, and kurtosis of the time series data, and also shows the Jarque-Bera test, which is a normality test. Jarque-Bera statistics show the normality of the series distribution has been rejected. Ljung-Box statistics shows significant autocorrelation in RV series which means RV has long term memory characteristic.

<table>
<thead>
<tr>
<th>Series</th>
<th>Q(5)</th>
<th>Q(10)</th>
<th>Q(15)</th>
<th>Q(20)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Return</td>
<td>8.306(0.1401)</td>
<td>14.707(0.1431)</td>
<td>17.443(0.2931)</td>
<td>19.830(0.4686)</td>
</tr>
<tr>
<td>RV</td>
<td>2352.009(0)</td>
<td>35558.792(0)</td>
<td>4363.946(0)</td>
<td>4830.773(0)</td>
</tr>
</tbody>
</table>

Table 1. Data Description

<table>
<thead>
<tr>
<th>Series</th>
<th>Mean</th>
<th>Sd</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>J-B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Return</td>
<td>0.000165</td>
<td>0.012010</td>
<td>-0.4832</td>
<td>3.87528</td>
<td>946.1795</td>
</tr>
<tr>
<td>RV</td>
<td>0.816720</td>
<td>0.353208</td>
<td>1.74806</td>
<td>5.12365</td>
<td>2288.684</td>
</tr>
</tbody>
</table>

4.2 Volatility Prediction

The rolling time window method was used for volatility prediction. First, I trained three single GARCH-type models which are GARCH, eGARCH, and tGARCH using rolling window of 252 days. Then volatility estimates from each GARCH model is used as the only input into LSTM to obtain volatility forecasts. After this, I generated a Hybrid model which combined four LSTM models to obtain final volatility estimates. The out-of-sample forecast performances are evaluated using four loss functions: MAE, MSE, RMSE and MAPE.
It can be seen from the following Table 2 that LSTM neural networks with RV as the only input gives the highest prediction accuracy. The result from this best LSTM-GARCH hybrid model was 0.2077(MAE), 0.0882(MSE), 0.2970(RMSE), 0.2378(MAPE), better than all other models.

Table 2. Model Performance

<table>
<thead>
<tr>
<th>Model</th>
<th>MAE</th>
<th>MSE</th>
<th>RMSE</th>
<th>MAPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>sGARCH</td>
<td>0.426390</td>
<td>0.283051</td>
<td>0.532025</td>
<td>0.562082</td>
</tr>
<tr>
<td>eGARCH</td>
<td>0.432255</td>
<td>0.265077</td>
<td>0.514856</td>
<td>0.573820</td>
</tr>
<tr>
<td>tGARCH</td>
<td>0.419351</td>
<td>0.231844</td>
<td>0.481502</td>
<td>0.567385</td>
</tr>
<tr>
<td>RV_LSTM</td>
<td>0.207696</td>
<td>0.088230</td>
<td>0.297035</td>
<td>0.237824</td>
</tr>
<tr>
<td>sGARCH_LSTM</td>
<td>0.362487</td>
<td>0.177257</td>
<td>0.421020</td>
<td>0.492438</td>
</tr>
<tr>
<td>eGARCH_LSTM</td>
<td>0.352927</td>
<td>0.169016</td>
<td>0.411116</td>
<td>0.476406</td>
</tr>
<tr>
<td>tGARCH_LSTM</td>
<td>0.356947</td>
<td>0.172733</td>
<td>0.415612</td>
<td>0.482356</td>
</tr>
<tr>
<td>Hybrid</td>
<td>0.308624</td>
<td>0.135770</td>
<td>0.368470</td>
<td>0.410260</td>
</tr>
</tbody>
</table>

Figure 2 shows the volatility predicted by the GARCH-type models versus the realized volatility. Figure 3 shows the volatility predicted by the LSTM models and Hybrid model.
4.3 Var Analysis

Value at Risk is a common measure of the risk of loss. It could estimate how much a set of investments might lose given a certain probability in a fixed time period such as a day. An important usage of VaR is risk management. It is defined as follows:

\[ P(r_{t+1} > VaR_{t+1}(\alpha)) = 1 - \alpha \]  \hfill (13)

\[ VaR_{t+1}(\alpha) = \mu + t_{\alpha}\sigma_{t+1} \]  \hfill (14)

In which \( \mu \) denote the mean of the return, \( t_{\alpha} \) denote the \( \alpha \) quantile of distribution of return time series, \( \sigma_{t+1} \) is obtained by the models mentioned in previous section.

In this study, VaR forecast is obtained by parametric method. Using the volatility forecast by RV-LSTM and Hybrid model to generate one step ahead of VaR. To improve the robustness of the experiment result, the confidence level \( 1 - \alpha \) was selected 90%, 95%, 99%, respectively. Figure 4 shows the predicted VaR versus return time-series. It can be seen from the figure,
hybrid model could make better use of the information from data to improve the accuracy of VaR forecasting.

Figure 4. VaR Prediction.

5 RESULTS

In this study, I proposed a hybrid model combing LSTM and GARCH-type models. The rolling window method is used in the experiments. Fixing the window size to be 22 trading days for the one-day ahead predictions of VaR. And according to the estimated VaR, I built a trading strategy to show the model performance. Negative VaR means the potential loss.

\[
\text{Flag}_{t+1} = \begin{cases} 
0, & \text{VaR}_{t+1}(\alpha) < -1 \\
1, & \text{VaR}_{t+1}(\alpha) \geq -1
\end{cases}
\]  

The function above is often called hitting series. It means if the VaR estimate at time $t+1$ is smaller than -1, the value of Flag at time $t+1$ equals to 0. In other words, if $\text{VaR}_{t+1}(\alpha) < -1$, the amount of funds I reserve will be insufficient to cover the potential loss, which means I need to sell the portfolio to avoid potential risk. According to this, I build a trading strategy and compare the performance with simple holding method. The results are shown in the following table and figures. Compared strategy performance with best single input RV-LSTM model, the cumulative return obtained by Hybrid model improved by 2.87, 1.97 and 2.60 under 90%, 95%, and 99% confidence level, respectively. It can be seen from both Table 3 and Figure 5 that the model improved the performance of the simple trading strategy in risk management significantly.
Table 3. Final Strategy Cumulative Return

<table>
<thead>
<tr>
<th>VaR</th>
<th>Hybrid Model</th>
<th>LSTM_RV</th>
<th>Improved by</th>
</tr>
</thead>
<tbody>
<tr>
<td>90%</td>
<td>3.6560</td>
<td>-6.8278</td>
<td>2.87</td>
</tr>
<tr>
<td>95%</td>
<td>4.6786</td>
<td>-4.5151</td>
<td>1.97</td>
</tr>
<tr>
<td>99%</td>
<td>1.5087</td>
<td>-2.4197</td>
<td>2.60</td>
</tr>
</tbody>
</table>

Figure 5. Cumulative Return of strategy.

6 CONCLUSION

This study proposed a hybrid model to combine several GARCH-type models with LSTM. It makes it possible to acquire various economic characteristic information. The magnitude of volatility shock and the persistence of volatility could be reflected in GARCH model. The persistence of volatility and the leverage effect could be reflected in the EGARCH model. Then I use information obtained in GARCH-type models as input into LSTM. LSTM could learn high-level temporal patterns in the time-series data by itself. Giving more information, the volatility pattern could be learned more efficiently, which could improve the prediction accuracy.
in the end. In order to prove this, I compared the performance of the hybrid model with single
GARCH-type model by testing these models’ performance on four different loss functions and
using them to predict the realized volatility of CSI 300 index data.

Finally, the hybrid model combing GARCH-type models and LSTM has great improvement on
prediction performance over single GARCH-type model. Compared to the single GARCH
model with best performance, the Hybrid model improved performance by 16%, 41%, 23% and
28% for MAE, MSE, RMSE, MAPE, respectively. Therefore, it could be confirmed that the
out-of-sample prediction error of the hybrid model is lowest for all the measures.

The empirical results in this study also showed that the hybrid model could significantly
improve the VaR prediction performance of the CSI 300 Index. By designing a simple trading
strategy according to VaR estimates, the cumulative return could be improved by 2.87, 1.97 and
2.60 under 90%, 95%, and 99% confidence level, respectively. Therefore, the research methods
and the conclusions of this study could provide the possibility for the further application of this
hybrid model in financial research.

With markets becoming more complex, to make a better volatility prediction in future research,
we should consider more diversified and dynamic information as indicator to predict volatility
such as some financial news in social website as inputs.

REFERENCES

provide accurate forecasts. International economic review, 885-905.
econometrics, 31(3), 307-327.
of China stock market. In 2015 IEEE international conference on big data (big data) (pp. 2823-2824).
IEEE.
and the volatility of the nominal excess return on stocks. The journal of finance, 48(5), 1779-1801.
1780.
integrating LSTM with multiple GARCH-type models. Expert Systems with Applications. 1;103:25-
37.
Manipulation Strategy Based on LSTM Networks for Stock Market Volatility. IEEE Access, 10, 34743-
34754.
Research on BIM Based EPC Project Lifecycle Information Management

Yuzhuo Zhang\textsuperscript{1,a}, Yuxuan Zhu\textsuperscript{1,b*}, Feng Guo\textsuperscript{1,c}, Xuhang Han\textsuperscript{1,d}

zhangyuzhuo@sjzu.edu.cn\textsuperscript{1}, *zhuyuxuanyx@foxmail.com\textsuperscript{b}, guofeng_fguo@163.com\textsuperscript{c}, xh_han@foxmail.com\textsuperscript{d}

School of Management, Shenyang Jianzhu University\textsuperscript{1}

Abstract: The construction mode and management mode of Chinese traditional construction industry are slow in informatization and digitization. This paper takes the information management of Engineering Procurement Construction (EPC) projects as the research object and Building Information Modeling (BIM) technology as the means to build a BIM cooperation platform in the information management of EPC projects, proving that the use of BIM technology in the whole life cycle information management of EPC projects can solve the problem of "information island" and promote the transmission and sharing of information in the whole life cycle of projects. To realize the dynamic monitoring of project safety information, and provide theoretical reference for realizing more efficient information management.

Keywords: EPC Projects, BIM Technology, Full Life Cycle, Information Management.

1 INTRODUCTION

In recent years, the problem of relative separation of each stage has emerged in our large complex engineering projects \cite{1}. There are barriers in the process of information data transmission in the whole life cycle of the project, which makes it difficult for managers to accurately grasp and effectively control the actual situation of the project. Combined with the advantages of EPC mode and BIM technology, the information management of collaborative projects in the whole life cycle can promote the information sharing and transmission between various stages and solve the problem of "information isolated island".

Cheng Xing et al. \cite{2} took actual projects as examples to study the application of BIM in each stage of EPC projects, and analyzed and summarized the advantages of BIM application in EPC projects. Guo Ziqi et al. \cite{3} put forward the method of applying BIM information model to the cost management of EPC projects, achieving the purpose of cost saving and providing a feasible idea for the combination of EPC project cost management and BIM technology. She Jianjun et al. \cite{4} proposed a BIM based knowledge integration management model for EPC projects.
projects, which orderly and systematized project knowledge. Amos Darko et al. [5], found through their research that BIM technology plays an important role in project risk management. To sum up, scholars have studied and explored the management of BIM technology in the aspects of cost, risk, and knowledge integration of EPC projects, but there are few researches on BIM technology in the whole life cycle information management of EPC projects. Therefore, the author takes the information management of EPC projects as the research carrier and BIM technology as the auxiliary tool to build a BIM based information management collaboration platform for EPC projects in the whole life cycle, to break the barriers of information transmission and promote the transmission and sharing of information in the whole life cycle of projects.

2 THE DEFINITION OF THE CONCEPT OF RESEARCH OBJECT

2.1 Conceptual Characteristics of EPC Projects

Engineering Procurement Construction (EPC) refers to a construction mode in which contractors’ contract for the whole process or several phases of a construction project's design, procurement, construction, and trial run according to the contract agreement and are fully responsible for the quality, schedule, and cost of the project [6]. And projects with this contracting process are called EPC projects.

EPC general contractors have greater freedom to play, and can carry out more effective supervision and management. In the EPC general contracting mode, the design, procurement, and construction can cross each other in the schedule according to the project requirements, so that all stages of the project construction can be effectively communicated and coordinated, to improve the construction efficiency of the project.

2.2 The Value of BIM Information Management

1) Implementation of associated modification: The software developed based on BIM can directly carry out 3D visual design, and provide the required data information such as model and plan for the subsequent stage. This greatly improved the shortcomings and deficiencies of the information sharing in the segmentation design in the CAD era, and reduced the occurrence of the failure to discover the design errors in time under the traditional mode, which would lead to the subsequent need to spend a lot of time and human and financial resources to remedy the situation. The BIM has the function of automatic coordination for modification. When the design is modified, the modified part can be synchronized to other associated content. It not only improves the design quality, saves people, money, and time, but also coordinates the information transfer and sharing between different specialties and different stages. It provides an important and reliable guarantee for the completion of the final goal [7].

2) Reduce repeated input of information: As the carrier of data, BIM can integrate different content data such as structural analysis and schedule management at each stage, so that data can be shared. In the process of schedule control, cost management, energy analysis and other work,
the required information can be directly extracted from the BIM database for research, and the results can be fed back to the database to prepare for the subsequent environment. Each link does not need to repeat the input and export of data, and will not cause data omission and loss. The information reuse rate of the whole life cycle of construction projects is improved.

3) The collaborative work of all participants is realized: All participants can obtain the required formal documents from the platform according to their own needs to meet the needs of collaborative work and project negotiation.

4) Improved degree of automation: the virtual construction of BIM can predict possible problems and dangers before construction, and then optimize the construction scheme and adjust the schedule based on this, which greatly improves the degree of automation in the construction process.

5) Support operation and maintenance activities: BIM database and the constructed information cooperation platform can be used to provide data protection for operation and maintenance stage, and provide comprehensive and perfect data support for building maintenance management software. Eliminate the phenomenon of data fragmentation when the information of design, construction and other stages is transmitted to the operation and maintenance stage.

6) Data interoperability: Data storage, exchange and sharing based on BIM can realize smooth interoperability among all links in the whole life cycle of the project and avoid the problem of “information isolated island”.

3 BUILDING AN INFORMATION MANAGEMENT COLLABORATION PLATFORM

3.1 Information Flow Between Project Participants

The responsibilities and roles of the participants in the EPC project lifecycle are different. Therefore, it is necessary to extract the required information and start the work. Then, the information generated by each participant is synchronized to the BIM database to coordinate the work of other participants [8]. The specific content is shown in the figure below, in which the light blue area is the direct participant of the project, and the other participants are the associated participants.

Figure 1. All parties involved in the whole life cycle of the construction project
There are different parties involved in each stage of the EPC project lifecycle to work together. Therefore, there are many participants and they communicate closely with each other. The flow of information between project participants is shown in the figure below.

![Figure 2. Information flow between key participants in the project](image)

### 3.2 The Construction Principle of Information Management Collaboration Platform

1) Real-time: The information collaboration platform should ensure that all project participants can work at any time and place, and are not limited by time, space, and other factors [9]. In addition, the information collaboration platform should ensure that participants can obtain the required information and data at any time to complete their own work. The construction of the platform should aim at strengthening information cooperation and sharing, strive to eliminate barriers to information transmission, and improve the quality and efficiency of information management. Ensure that the platform synchronizes and updates the real-time status of construction projects at any time. Solve problems in time, ensure and coordinate the smooth work of project participants.

2) Fluency: The information management cooperation platform should ensure the high efficiency of information exchange and transmission among participants and between different stages. Therefore, the collaborative platform needs to provide data exchange functions between IFCs to assist BIM in building a complete data system and ensure the smoothness of information transmission.

3) Security: Due to the characteristics of multi-subject construction projects, the construction scope of each subject is different, so it is necessary to establish security measures to protect the interests of all parties from the harm of virtual network risks. It is necessary to build the information cooperation platform with security, establish access rights, improve security measures to ensure the information security in the whole life cycle of project construction.

### 3.3 The Content of Information Management Collaboration Platform

1) Establish BIM database. It is beneficial to realize the information integrated management of EPC projects and meet the data acquisition requirements of EPC teams and participants. Participants can extract data information from the database according to their own needs, and upload the data and results generated by the work to the BIM database after completing the work content they are responsible for.
2) It is equipped with BIM software to satisfy the information management and application of all participants in the whole life cycle. In order to satisfy the information management and application of EPC project participants, the software equipped should meet the application in the process of project implementation. Information interaction between BIM software is the basis of information collaboration platform.

3) Information management of EPC project participants. Each participant can have different permissions according to their own needs to obtain information and data from the collaboration platform, and the management authority of the EPC general contractor should be the highest level to ensure that the project information is fully mastered.

3.4 Establish BIM Model in EPC Project

The establishment of BIM model for EPC projects is the core of applying BIM to the whole life cycle information management of EPC projects. The integration and interaction of information in the project is also realized based on BIM data. Extract the information in the basic data layer to build the BIM information model. BIM models and projects are associated with the collaboration platform to ensure the transmission and sharing of information data in the whole life cycle of EPC projects.

4 INFORMATION MANAGEMENT FOR THE WHOLE LIFE CYCLE OF EPC PROJECTS

4.1 Information Interaction Management Between Three Stages of The EPC

The three stages of the EPC include the engineering phase, procurement phase and construction phase. The EPC project has many participants, which increases the difficulty of coordination for the general contractor \(^{[10]}\). Therefore, an information management framework for EPC projects based on BIM should be established to give full play to the advantages of BIM, improve the information management ability and help the integrated management of design, procurement, and construction of EPC projects.

![Figure 3. Information interaction between EPC three stages](image-url)
4.2 Information Interaction Management of EPC Phase 3 and Operation Phase

The information of the project operation management phase also needs the support of the information of the EPC phase, to carry out regular maintenance and management of the project. The efficient interaction of EPC phase 3 and operation phase information can ensure that the service performance of building facilities is always maintained in a good state, extending the service life, and realizing sustainable application.

4.3 EPC Project Life Cycle Safety Monitoring Information Management

1) Security detection process: effectively distinguishing and identifying security vulnerabilities during the construction period of the project is the key to ensure the safety of the project. Through the information collaboration platform built by BIM technology, project participants can simulate the whole process of EPC project before the project starts. Identify possible security problems and formulate security prevention plans. The following figure shows the security management system built based on BIM technology.

![Figure 4. Information interaction between EPC three stages](image)
Figure 5. EPC Phase 3 and operation phase information interaction

Project managers can accurately identify construction risks, formulate scientific prevention plans, and implement integrated construction plans. With the real-time performance of BIM technology, security risks can be dynamically identified to reduce unnecessary losses.

2) Safety inspection of virtual construction: the information cooperation platform built based on BIM can carry out visual supervision on the entire construction process. The participants can timely understand the requirements of the work and their own responsibilities, and it is convenient for managers to grasp the use status of materials on site. The information data of each stage and participants are timely fed back to the BIM collaboration platform, which is convenient for managers to efficiently evaluate construction schemes, find problems and loopholes, analyses, and solve problems. Truly achieve the security of information science control and management.

5 CONCLUSIONS

BIM technology helps EPC projects carry out information management, breaks the information barriers of information data transmission at all stages of EPC project life cycle, and solves the problem of "information island". BIM based EPC project collaboration platform can promote
information transfer and sharing at all stages of the project. It facilitates the reliable and convenient information exchange in the whole life cycle of EPC projects, realizes the cooperation and information integration between each stage of the project and each participant, improves the efficiency of information management, and makes the project obtain considerable benefits. In addition, the safety management system based on BIM technology can accurately identify project risks and prevent and avoid them, which greatly reduces unnecessary project losses and reduces construction costs.

ACKNOWLEDGMENTS This study was supported by Doctoral Scientific Research Foundation, China (2019-BS-197); Liaoning Basic Research Project, China (lnjc202007); Liaoning Provincial Colleges and Universities' Innovative Talents Support Plan (LR2020005).

REFERENCES


Application Design of Road, Bridge and Tunnel Engineering Information Management Based on Blockchain

Ping Zhang\textsuperscript{1,a}, Mingchao Liao\textsuperscript{1,b}

\textsuperscript{1}1140766075@qq.com, 411781160@qq.com

Abstract: In the process of information management of road, bridge and tunnel projects in the field of infrastructure, the traditional centralized management method has problems such as difficulty in data supervision, difficulty in data traceability, and easy tampering of information. To this end, combined with the blockchain technology, a blockchain application architecture based on the underlying architecture of the alliance chain and the middleware platform is proposed—the road, bridge and tunnel project information chain architecture. The test results show that the application can not only effectively ensure the security of road, bridge and tunnel engineering information, but also has high performance, and has certain application value in the management of road, bridge and tunnel engineering information.

Keywords: Engineering Information Management, Blockchain, Application Architecture.

1 INTRODUCTION

With the rapid growth of the scale of China's transportation infrastructure construction, the rapid improvement of the level of industrialized manufacturing, and the continuous advancement of digital construction, combined with the urgent needs of China's road and bridge tunnel industry upgrading and transformation, how to further improve the overall level of informatization of the road and tunnel industry is an important task \cite{1} in front of us, which coincides with the development of new infrastructure \cite{2}. To this end, researchers have designed distributed information management applications for road, bridge and tunnel engineering, which greatly improves the efficiency of information management, but there are still problems such as easy data tampering, difficult supervision of users, and untrustworthy traceability. The birth of blockchain \cite{3} technology provides a solution to the problem of data privacy, security and integrity of road and bridge information management. A. Lanko, N \cite{4} will play the role of blockchain technology in the safety traceability of bridge construction, electronically record the key information of each link through RFID technology, record it in the bridge database based on blockchain in real time, and form tamper-resistant information after authentication, realize the tracking and traceability management of bridge engineering safety, and use blockchain nodes to represent all transaction parties in the bridge engineering construction process, and any
information processing behavior of transaction parties on the blockchain is transparent. So as to achieve effective supervision of all parties involved in bridge engineering.

In order to solve the above problems, this paper constructs a road-bridge engineering information management application based on FISCO BCOS blockchain, which integrates the road, bridge and tunnel through detailed architecture design, and uses smart contract technology and test its performance.

2 RELATED THEORIES

2.1 Smart Contracts and Blockchain

The concept of smart contracts, proposed by Nicosab in 1995, is designed to digitally define a code container that reflects a real-world contract, and when the participants meet the required conditions of the contract and execute it, the computer will automatically execute the contract agreement within the code container. Due to the technical conditions at the time, smart contracts were not used. Blockchain originated from the Bitcoin system, which is essentially a peer-to-peer distributed system with a unique consensus protocol and incentive mechanism that guarantees the decentralization, transparency, traceability, trustworthiness and immutability of system operations. Bitcoin is also known as Blockchain 1.0. The combination of smart contracts and blockchain technology gave birth to Ethereum with the goal of solving the problem of mutual trust in the financial field, which is called Blockchain 2.0. In order to improve the universality of blockchain technology, researchers propose that blockchain should go beyond the financial field, provide decentralized solutions for various industries, and build an intelligent Internet of Things era called blockchain 3.0. This article uses blockchain 3.0 technology to build blockchain applications.

2.2 Related Study

Blockchain technology as an emerging technology has been deeply discussed and tried to be applied to all walks of life, among which Liu Tianyu and others focused on analyzing the application scenarios of blockchain technology and the current problems, and looked forward to the future technology trends, which clearly proposed that the combination of blockchain technology and transportation infrastructure is one of the major topics in the development of blockchain technology, and the application of blockchain technology can provide solid technical support for the future construction of road, bridge and tunnel infrastructure. Caixiang Fan tested and analyzed the performance of the current mainstream blockchain in different application scenarios, among which the performance of the consortium chain is much better than the performance of the public chain, FISCO-BCOS in the consortium chain is the least delayed, and the transaction throughput is better than that of Hyperledger Fabric, which is also a consortium chain. Literature et al., as the underlying developers of FISCO Alliance Chain, further discussed the application form based on FISCO BCOS based on practical application landing examples, and summarized its layout and practice in technology, application and ecological construction: FISCO is suitable for providing information technology support for intelligent transportation infrastructure construction.
In order to solve the problem of modern traffic information management, researchers at home and abroad have used a variety of advanced technologies. Among them, literature [11] constructed a BIM bridge model, which visually displays the progress of bridge construction through the model, and compares the planned progress with the actual progress.

The information management technology of the construction process of the steel precast components of the main beam of the bridge was developed to realize the information sharing and construction control in the construction process. However, the study only uses general database technology and document management technology to build the data layer, and the entire system lacks the protection of data security, and there is a problem that the record data is maliciously tampered with, resulting in untrustworthy system management. Literature [12] developed a blockchain-based quality information management framework covering Hyperledger Fabric-based architecture and a series of solutions, and experiments verify that the framework can decentralize the management of quality information to achieve consistent and secure quality information management. However, the underlying blockchain technology used by the framework has the problem of weak performance and high latency [9].

Based on the above analysis, blockchain technology can be applied to provide information technology for the information management of transportation facilities such as roads, bridges and tunnels, and blockchain technology is more secure than traditional information management and data storage technology.

3 ARCHITECTURAL DESIGN

3.1 Operate Architecture

The architecture design [13] is divided into two parts: architecture construction and architecture operation, and the architecture construction process is shown in Figure 1: first, the receiving party, regulatory agency and each participant deploy the underlying network nodes representing themselves locally, and connect with each other through the network interface to form the underlying node network of the alliance chain. Then, the acceptor builds a middleware platform locally, connects with the application service layer through the network interface to form a web network service, connects with the bottom layer of the consortium chain, and creates a database as a storage location for block information. Finally, the receiving party accesses the middleware platform through the web network service, deploys the smart contract, and generates the contract interface. The construction process is shown in Figure 1 alphabetical order: first the participants access the web network service, and the web service determines whether the participant has the right to use the contract interface, if it does not have permission, the access request is rejected, and if it has permission, the participant can call the contract interface to execute the smart contract function, and record the road and bridge tunnel project information and the transaction information between the participants in the bottom layer of the alliance chain.
### 3.2 Node Group Architecture

In order to adapt to the requirements of real business scenarios, FISCO BCOS adopts a group architecture to support blockchain nodes to start multiple groups, and the transaction processing, data storage, and block consensus between groups are isolated from each other, which not only ensures the privacy of blockchain applications \[14\], but also reduces the complexity of O&M. In the multi-group architecture, the consortium chain nodes are divided into free nodes and group nodes, and each node in the group is divided into consensus nodes and observation nodes, as shown in Figure 2 below.
Figure 2 has three parts of repeated graphs, each of which represents a node group, which in turn can be divided into roads, bridges, or tunnels. Each group contains all the nodes of the node layer in Figure 1, which are owned by the participants and regulatory agencies in the road and bridge tunnel construction process, and in particular, the acceptance nodes and a group of observation nodes in each group form a group 4. The construction logic of the node group architecture is: group 1, group 2 and group 3 each represent an engineering project, the transportation, construction, manufacturing, and maintenance of the project are held by the corresponding participants, when the engineering information needs to be recorded, these nodes and the acceptance nodes owned by the manager use the PBFT consensus algorithm to reach consensus, and after the consensus is successful and the observation node held by the regulator is confirmed, the block containing engineering information will be on the chain. Group 4 is a special group, the consensus process is for the three-party project managers to reach a consensus using the PBFT consensus algorithm, and the information block is uploaded to the chain after being confirmed by the regulator.

4 APPLICATION PERFORMANCE TESTING

JMeter is used for performance testing: JMeter relies on the Java environment, and after configuring the JDK, it simulates the scenario of multiple users calling the web server interface concurrently by setting a thread group with a fixed request time. Use this tool to test the contract deployment interface and transaction execution interface of the application, and obtain the interface call error rate ERR, transaction throughput per second TPS, and average transaction response time RTT, and obtained conclusions based on these data.
JMeter was used to stress test five sets of different numbers of threads for the contract deployment interface and the transaction execution interface. First, set the number of thread groups to 500, 1000, 2000, and 4000, with a request interval of 1 second and a maximum thread duration of 900 seconds. Then set the sampler to HTTP request, add the server IP, port number, request mode, request parameters and other information, and save it as a JMX file. Finally, run the jemeter command in the command window to run the JMX file, and the test results are shown in Table 1 below.

<table>
<thead>
<tr>
<th>sum</th>
<th>Contract interface</th>
<th>Trade execution interface</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ERR/%</td>
<td>TPS</td>
</tr>
<tr>
<td>500</td>
<td>0</td>
<td>5.3</td>
</tr>
<tr>
<td>1000</td>
<td>0</td>
<td>6.9</td>
</tr>
<tr>
<td>2000</td>
<td>0.15</td>
<td>6.6</td>
</tr>
<tr>
<td>3000</td>
<td>4.03</td>
<td>7.5</td>
</tr>
<tr>
<td>4000</td>
<td>10.20</td>
<td>6.1</td>
</tr>
</tbody>
</table>

5 CONCLUSION

Aiming at the problems of difficult data supervision, data traceability and easy tampering of information in road and bridge tunnel engineering, this paper designs an information management application based on the FIS-CO BCOS alliance chain. The architecture of the application organically integrates all participants in the road and bridge tunnel engineering construction process, and uses secure and loophole smart contracts to construct application services, which ensures the security of engineering information and realizes the digital management of road and bridge tunnel engineering information. It can be seen from the experimental results that application performance can achieve the effect required by the actual application. The next research of the application is to dynamically increase the number of blockchain nodes and expansion groups online, accept the data incoming from physical devices, and comprehensively manage the collected information.

ACKNOWLEDGMENTS This work was financially supported by the Science and Technology Project of Hubei Provincial Department of Transportation (2020-2-6-1).

REFERENCES


Research on the Mediation Model and Empirical Analysis for the Influence of Digital Investment on Retail Business Performance

Xing XING¹, Qinglie WU²,*

Corresponding author: wql@seu.edu.cn²*

School of Economics and Management, Southeast University, Nanjing, 211189, China¹
Jiangsu Academy of Smart Industries and Digitalization, Nanjing, 210031, China²

Abstract: The purpose of this paper is to systematically analyze the influence of digital input on the business process and business performance of retail businesses, and to establish a model for empirical verification with quantitative methods. Based on the data of A-share retail listed companies from 2010 to 2020, this paper establishes a mediation model and verifies the impact and mechanism of digital investment on the performance of retail businesses. The empirical results show that the digital investment of businesses can positively affect the performance of businesses. From the perspective of influencing mechanism, digital investment affects business performance through three paths: digital operation and management, digital innovation of retail, and business innovation outside retail. In the conclusion, how to understand the role of digital technology in retail operation, how to use the intermediary model to guide businesses to make targeted investment, and how to strengthen the determination of businesses to digital investment and transformation are further explained.

Keywords: Digital Transformation, Digital Investment, Mediation Model, Empirical Analysis, Business Performance.

1 INTRODUCTION

Looking back at the previous research on digital transformation of retail industry, scholars mostly qualitatively study the mechanism and connotation of digital transformation through cases [4, 13]. The digital transformation of the retail industry is essentially the reconstruction of people, goods and market elements [9], with consumer demand as the core and digital technology as the driving force, reengineering business processes.

However, there is little quantitative research on the performance of retail businesses by digitalization at present. Dong (2022) based on the data of 76 listed retail businesses and the perspective of channel concentration, verified that the digital transformation of retail businesses is conducive to improving their operational performance and enhancing their profitability and capital planning ability; Huang (2022) found that digital transformation has a positive impact on business efficiency by expanding the market scale and improving management efficiency, and the excessive scale will inhibit the promotion of retail businesses' digital transformation on business efficiency. Similar to “IT productivity paradox”, some scholars found that the business...
efficiency of dual-channel retail businesses after O2O transformation did not significantly improve [15].

Does digital investment increase the cost of the business and thus reduce its performance? Or through other ways to improve corporate performance? Therefore, clarifying the impact of digital investment on the performance of retail businesses in the process of digital transformation, as well as the mechanism and path behind it, and making a reasonable investment plan according to the characteristics of the retail industry will help businesses solve the problems of "dare not turn" and "won't turn", and help them formulate a reasonable digital strategy.

2 THEORETICAL ANALYSIS AND RESEARCH HYPOTHESIS

2.1 The Influence of Digital Investment on the Retailing Performance

The process of digital transformation of businesses is the reconstruction of business, process and organization by digital technology [12], and its ultimate goal is to improve the competitiveness and profitability of businesses. First of all, digital transformation is a process of applying digital technology to organizational change [3]. Digital technology has broken through the digital barrier between business departments and even between businesses, and through the establishment of collaborative system and resource integration, the organizational structure of businesses is more agile and flexible, thus improving the internal operational efficiency and organizational performance of businesses. Secondly, digital technology has significantly improved the information processing capability of businesses [11].

In the process of digital transformation, retail businesses will generate a large amount of transaction data. The application of big data, cloud computing and other technologies makes these previously unknown data "live" and become a brand-new production factor of businesses, and tap its value to help improve the efficiency of decision-making, production, circulation, marketing and other processes of businesses. In the digital era of changing consumer demand, fiercer competition among businesses, shorter product cycle, and full of complexity and uncertainty in the market, digital technology can help businesses accurately grasp the demand and gain insight into the market trend, so as to revise and formulate appropriate strategies in time.

According to Accenture's Digital Transformation Index of China Businesses, from 2015 to 2017, the compound growth rate of digital transformation leader's operating income was as high as 14.3%, while that of other businesses was only 2.6%. In 2017, the former's sales profit margin reached 12.7%, while the latter was only 5.2%.

So, the hypothesis is proposed in this paper:

**H1**: Digital investment has a positive impact on the performance of retail businesses.

2.2 Mediating Mechanism by Which Digital Investment Affects Retailing Performance

The essence of the process of digital transformation is the digital transformation of management mode [7]. Digital management is a management method that utilizes digital technology to
quantify management objects and behaviors and realize functions such as planning, organization, production, sales, service, innovation, etc., and introduces digital technology into business operation and management to promote the systematic reshaping of business organizational structure, information flow, management methods and operation mechanism. The development and popularization of digitalization has enhanced the insight of businesses into all aspects of operation and management, which can better make operation and management decisions and improve operation efficiency.

The improvement of digital technology on the operation and management efficiency of retail businesses is mainly reflected in the following two aspects. On the one hand, the digital transformation makes the organizational structure tend to be networked and flat \[10\]. The application of digital technology speeds up the circulation of information elements in businesses, and managers can quickly perceive the changes in the business and the market and take prompt measures to deal with them. In addition, digital technology connects the "digital islands" of different departments of the business, strengthens data sharing among different departments, promotes communication and learning within the organization, stimulates the innovation of employees' production management, and empowers total factor production. For example, through big data analysis and mining of employees' career information, businesses build a big data platform for human resources, which brings employee recruitment and talent training into the quantitative category and provides reliable basis for management decisions such as performance assessment, employee motivation and employee promotion.

On the other hand, digitalization and related technologies have played an important role in the innovation of supply chain management of businesses \[11\]. For example, block chain technology is essentially a distributed and immutable database. As a distributed ledger technology to ensure transparency, traceability and security, it can play a huge role in product traceability and strengthening trust between businesses. Big data analysis and cloud computing technology solve supply management, demand forecasting, inventory control and transportation route design in operation by analyzing a large amount of supply chain operation data. By incorporating digital technologies, it reduces out-of-stock rates, reduces purchase lead times, reduces inventory levels throughout the supply chain and increases corporate profits.

Therefore, the hypothesis is put forward in this paper:

**H2:** Digital investment positively affects the performance of retail businesses by improving the efficiency of digital operation and management.

In the process of digital transformation, digital technology promotes the digital innovation and development of retail business model. Under the background of digital economy, traditional retail businesses are faced with more and more personalized consumer demand, the impact of e-commerce businesses and many other pressures. The traditional business model is difficult to guarantee the profit source of businesses, and the value growth has shifted to online and offline service activities.

Therefore, businesses use digital technology to optimize the entire industrial value chain, integrate and continuously improve users' online and offline shopping and service experiences, and maintain and enhance their competitive advantages. The digital innovation of retail is a new retail model that takes consumers as the core, digital technology as the driving force, realizes online and offline integration through intelligent infrastructure, reconstructs retail scenes, and
realizes full scene, full customer base, full data, full experience, and omni-channel integration. New retail formats generated by digital innovation of businesses emerge in endlessly: Super Species, Box Horse Fresh Life, Bailian RISO, Tianhong SP@CE, Easy Carrefour and other new retail formats appear and develop rapidly.

Compared with traditional retailing, the new retailing pays more attention to scene marketing and customer experience, and relies on the Internet and big data to deeply integrate online and offline, so as to realize the sharing of data information such as products, members, transactions and marketing, and provide consumers with the ultimate experience of cross-channel and seamless. For example, retail plus Internet of Things technology collects customers' movement track, attention change and customer flow information by using various sensors placed in the store for store operation and precision marketing; With retail and virtual reality technology, the visual display of goods will be upgraded from two-dimensional to three-dimensional. AR will be used to help customers experience the 3D layout effect of goods before purchase, and VR will be used to help customers test the product effect. Digital twin combined with AR/VR technology will make the interactive experience in the real sense become reality.

Therefore, the hypothesis is put forward:

**H3:** Digital investment positively affects the performance of retail businesses through digital innovation of retail.

At the same time, the digital transformation of businesses gives them more value-added opportunities in the value chain. For example, Yonghui Supermarket uses the massive consumer data accumulated in the process of business digitalization, analyzes the consumer behavior characteristics behind the data to judge its credit, and issues consumer loans to form a "cloud gold" financial business. Huadu businesses use consumer data to develop digital marketing services, e-commerce operation services and customized product marketing services. Hongqi Chain uses big data to develop IMP integrated marketing platform, and successfully builds a new model of community retail. Generally speaking, the more digital technology businesses invest, the better new models and new businesses will be developed, and the better business performance will be. Therefore, the hypothesis is put forward:

**H4:** Digital investment positively affects the performance of retail businesses through business innovation outside retail.

To sum up, digital transformation affects the performance of retail businesses by digital operation and management, digital innovation of retail, and business innovation outside retailing. And the theoretical model of this paper is established, as shown in Figure 1.
3 SAMPLE, VARIABLES, AND MODELING

3.1 Sample Selection

In 2010, with the rise of mobile Internet, new formats are constantly emerging, and netizens have an unprecedented impact on the social economy. The digital economy has entered a new stage, and the physical retail businesses in China have started the practice of digital transformation. Considering the listing time of domestic retail businesses and the impact of Covid-19 epidemic, the research time interval is locked in 2010-2020. Secondly, the retail industry includes many sub-industries. In order to ensure the comparability of digital transformation, business model and profit model, 30 major supermarkets listed on A-share market are selected as the research objects. The data of the research comes from the national CSMAR database and the annual reports of listed companies.

3.2 Variable Definition

- **Explanatory variables:** Digital investment.

According to the characteristics of retail businesses, this paper takes the digital investment of software, sales network, ERP and other digital technologies in the intangible assets of businesses as the digital investment of businesses, and measures the digital investment level by the ratio of digital investment to operating income.

- **Mediator variable:** Digital operation and management.

In this paper, the related word frequency of text mining, business management fee rate and inventory turnover rate are used to measure the digital operation and management of businesses. The higher the frequency, the higher the application of digital technology. The lower the management fee rate of an business, the higher the inventory turnover rate, indicating the higher the operation and management efficiency of the business. Therefore, the related word frequency * inventory turnover rate/management fee rate is used to measure the digital operation and management of businesses.
First, build a digital technology "thesaurus" shown in Table 1. Use jieba package in python software to segment the listed annual reports of selected companies, and learn from the research of scholars such as Wu (2021) and Zhao (2021) to form digital technology "thesaurus".

Secondly, use python to count the frequency of keywords in the "thesaurus" in the annual report. Finally, the result is calculated by the formula\[ M = \frac{m_i \cdot \text{inventory}_i}{\text{cost}_i} \] (word frequency, inventory turnover rate and management fee rate have been normalized in order to unify the dimensions).\[ m_i \] represents the number of digital technology keywords of \( i \) business in \( t \) year, \[ \text{inventory}_i \] represents the inventory turnover rate of \( i \) business in \( t \) year, and \[ \text{cost}_i \] represents management fee rate of \( i \) business in \( t \) year. The larger the \( M \), the higher the degree of digital operation and management.

- Mediator variable: Digital innovation of retail.

Under the background of digital transformation, retail digital innovation has become an important development strategy of businesses, and such important information will be elaborated in the annual report. Therefore, it is scientific and feasible to use text mining technology to count the keyword word frequency (normalized word frequency) about retail digital innovation in annual reports of businesses to measure the degree of retail digital innovation.

- Mediator variable: Business innovation outside retailing.

This paper adopts the growth rate of other business income as an index to measure the innovation outside retail business. Other business income of retail businesses includes financial services, digital operation services, advertising services, etc. The greater the growth rate of other business income, the better the innovation of businesses' off-retail business.


In this paper, referring to the research of Yi, Wu and other scholars (2021), the return on assets excluding financial income is used to measure business performance. Specifically, \[ R = \frac{\text{operating profit} - \text{investment income} - \text{fair value change income} + \text{investment income for joint ventures and joint ventures}}{\text{total assets}} \]

- Control variables.

Using previous scholars' relevant research for reference, the variables that may affect the innovation performance of businesses are selected as control variables. The variables and their symbols are shown in Table 2.

<table>
<thead>
<tr>
<th>Table 1: Digital technology keywords.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digital technologies related to digital operation management</td>
</tr>
</tbody>
</table>
Digital technologies related to digital innovation of retailing

Instant consumption, unmanned technology, virtual reality, omni-channel, live broadcast, mini program, online and offline, Retail format, experiential retail, scene retail, network collaboration, retail ecology

Table 2: Variables’ name and symbol.

<table>
<thead>
<tr>
<th>The variable</th>
<th>Variable symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explanatory variables</td>
<td>Digitization input D</td>
</tr>
<tr>
<td>Mediator variables</td>
<td>Digital operations and management M</td>
</tr>
<tr>
<td></td>
<td>Digital innovation of retail N</td>
</tr>
<tr>
<td></td>
<td>Business innovation outside retail B</td>
</tr>
<tr>
<td>The variable being explained</td>
<td>Business performance R</td>
</tr>
<tr>
<td>Control variables</td>
<td>Company age age</td>
</tr>
<tr>
<td></td>
<td>Cash ratio cash</td>
</tr>
<tr>
<td></td>
<td>Gearing ratio liability</td>
</tr>
<tr>
<td></td>
<td>Total asset scale</td>
</tr>
<tr>
<td></td>
<td>Year fixed effect year</td>
</tr>
</tbody>
</table>

3.3 Model Setting

Based on Wen (2004) practice, this paper adopts the three-step method to construct the mediator effect test model. Firstly, it examines the direct impact of digital investment on retail businesses.

\[
R_{it} = \omega_0 + \omega_1 D_{it} + \sum \omega_k Control_{it} + \epsilon_{it} \tag{1}
\]

Among them, \(\omega_0\) is a constant term, subscript \(i\) and \(t\) represent business and year respectively, \(Control_{it}\) represents control variables, and \(\epsilon_{it}\) are random error terms.

Secondly, test the influence of digital input on mediator variables.

\[
M_{it} = \alpha_0 + \alpha_1 D_{it} + \sum \alpha_k Control_{it} + \epsilon_{it} \tag{2}
\]

\[
N_{it} = \beta_0 + \beta_1 D_{it} + \sum \beta_k Control_{it} + \epsilon_{it} \tag{3}
\]

\[
B_{it} = \theta_0 + \theta_1 D_{it} + \sum \theta_k Control_{it} + \epsilon_{it} \tag{4}
\]

Finally, the mediator effect is tested. According to the previous hypothesis, digital investment can improve the performance of businesses by influencing mediator variables, so formula (2), (3) and (4) can be brought into formula (1).

\[
R_{it} = \gamma_0 + \gamma_1 D_{it} + \gamma_2 M_{it} + \gamma_3 N_{it} + \gamma_4 B_{it} + \sum \gamma_k Control_{it} + \epsilon_{it} \tag{5}
\]
4 EMPIRICAL RESULTS AND ANALYSIS

4.1 Descriptive Statistics

Table 3 is the descriptive statistical results of all variables in this paper. From the perspective of digital input level, the min is 0, the max is 0.095, and the SD is 0.025, which shows that the sample businesses have a big gap in digital investment. And the average value of 0.022 is greater than the median value of 0.012, and the data distribution is to the right, which shows that only a few businesses have a relatively high proportion of digital investment, and most businesses have insufficient digital investment.

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>AVG</th>
<th>SD</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digitization input</td>
<td>287</td>
<td>0.00</td>
<td>0.095</td>
<td>0.022</td>
<td>0.025</td>
<td>0.012</td>
</tr>
<tr>
<td>Digital operations management</td>
<td>286</td>
<td>0.00</td>
<td>4.184</td>
<td>0.364</td>
<td>0.540</td>
<td>0.186</td>
</tr>
<tr>
<td>Digital innovation of retail</td>
<td>287</td>
<td>0</td>
<td>1</td>
<td>0.027</td>
<td>0.064</td>
<td>0.013</td>
</tr>
<tr>
<td>Business innovation outside retail</td>
<td>276</td>
<td>-0.970</td>
<td>4.679</td>
<td>0.131</td>
<td>0.464</td>
<td>0.105</td>
</tr>
<tr>
<td>Business performance</td>
<td>287</td>
<td>-0.243</td>
<td>0.187</td>
<td>0.033</td>
<td>0.039</td>
<td>0.036</td>
</tr>
<tr>
<td>Company age</td>
<td>287</td>
<td>3.00</td>
<td>39.000</td>
<td>19.404</td>
<td>6.738</td>
<td>20.000</td>
</tr>
<tr>
<td>Cash ratio</td>
<td>287</td>
<td>0.026</td>
<td>2.546</td>
<td>0.527</td>
<td>0.465</td>
<td>0.376</td>
</tr>
<tr>
<td>Gearing ratio</td>
<td>287</td>
<td>0.146</td>
<td>0.943</td>
<td>0.578</td>
<td>0.157</td>
<td>0.589</td>
</tr>
<tr>
<td>Total assets</td>
<td>287</td>
<td>1.089</td>
<td>56.158</td>
<td>9.018</td>
<td>7.265</td>
<td>7.230</td>
</tr>
</tbody>
</table>

Note: Total asset unit: billion yuan

Similarly, it indicates that the sample businesses have great differences in performance and business innovation outside retail. In terms of digital operation management, similar to digital investment, the digital operation management level of a few businesses is higher, but the overall level is lower. In the aspect of digital innovation of retail, the innovation level of different companies fluctuates greatly between different years. In addition, there are significant differences in the control variables, and controlling them is effective for the study of firm performance.

4.2 Regression Analysis of the Influence of Digital Investment on Business Performance

Table 4 shows the basic regression results of the impact of digital investment on the performance of retail businesses. Column (1) shows that the regression is carried out by the ordinary least squares method (OLS) without adding control variables, and the regression coefficient value of digital investment is 0.345, but the result is not significant. Column (2) is the regression result with the addition of control variables, and the digital investment coefficient is significantly positive at the level of 0.01, which means that digital investment will have a significant positive impact on the performance of retail businesses, and R2 has been significantly improved, indicating that the degree of regression interpretation is further enhanced. In column (3), the
year fixing effect is added, and the results show that the significance and R² are obviously improved. To sum up, digital investment has a positive impact on the performance of retail businesses, and it is necessary and effective to add control variables and year fixed effect, so H1 has been verified.

### Table 4: Results of benchmark regression analysis of digital investment on the business performance.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.025*</td>
<td>0.052**</td>
<td>0.054**</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.015)</td>
<td>(0.013)</td>
</tr>
<tr>
<td>Digitization</td>
<td>0.345</td>
<td>0.382**</td>
<td>0.410**</td>
</tr>
<tr>
<td>investment</td>
<td>(0.093)</td>
<td>(0.084)</td>
<td>(0.731)</td>
</tr>
<tr>
<td>Company age</td>
<td>0.001</td>
<td>0.002*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.002)</td>
<td></td>
</tr>
<tr>
<td>Cash ratio</td>
<td>0.016*</td>
<td>0.004*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.006)</td>
<td></td>
</tr>
<tr>
<td>Gearing ratio</td>
<td>-0.076**</td>
<td>-0.133**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td>(0.018)</td>
<td></td>
</tr>
<tr>
<td>Total assets</td>
<td>0.000</td>
<td>0.001*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td></td>
</tr>
<tr>
<td>year</td>
<td>Uncontrol</td>
<td>Uncontrol</td>
<td>Control</td>
</tr>
<tr>
<td>N</td>
<td>287</td>
<td>287</td>
<td>287</td>
</tr>
<tr>
<td>R²</td>
<td>0.046</td>
<td>0.246</td>
<td>0.469</td>
</tr>
</tbody>
</table>

Note: * p<0.05 ** p<0.01 Standard error in parentheses

#### 4.3 Analysis of Mediating Mechanism

Table 5 shows the mediator role of digital operation and management, digital retail innovation and business innovation outside retail in the impact of digital investment on the performance of retail businesses. Column (1) shows the direct impact of digital investment on the performance of retail businesses, and digital investment has a significant role in promoting the performance of retail businesses. Columns (2) and (5) are the test results of the mediator role of digital operation and management. In column (2), the coefficient of digital investment is significantly positive at the level of 0.05, indicating that digital investment has a positive role in promoting the digital operation and management of retail businesses. Column (5) shows that digital operation and management can significantly improve the performance of retail businesses. Combined with the data in columns (2) and (5), it shows that digital operation and management has played a partial mediator role, and retail businesses have improved their digital operation and management level through digital investment, thus improving their performance. Generally, businesses with large digital investment will pay more and more attention to digital management of businesses, and apply the invested digital technology to the reform and upgrading of business management and supply chain management. H2 has been verified.
Table 5: Results of mediator effect test.

<table>
<thead>
<tr>
<th></th>
<th>Business performance</th>
<th>Digital operations management</th>
<th>Digital innovation of retail</th>
<th>Business innovation outside retail</th>
<th>Business performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.052** (0.013)</td>
<td>0.232 (0.239)</td>
<td>0.043 (0.027)</td>
<td>0.705* (0.277)</td>
<td>0.043** (0.013)</td>
</tr>
<tr>
<td>Digitization investment</td>
<td>0.399** (0.075)</td>
<td>3.155* (1.354)</td>
<td>0.314* (0.157)</td>
<td>6.084* (1.568)</td>
<td>0.318* (0.077)</td>
</tr>
<tr>
<td>Digital operations management</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Digital innovation of retail</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Business innovation outside retail</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year fixed effect</td>
<td>control</td>
<td>control</td>
<td>control</td>
<td>control</td>
<td>control</td>
</tr>
<tr>
<td>Other control variables</td>
<td>control</td>
<td>control</td>
<td>control</td>
<td>control</td>
<td>control</td>
</tr>
<tr>
<td>N</td>
<td>275</td>
<td>275</td>
<td>275</td>
<td>275</td>
<td>275</td>
</tr>
<tr>
<td>R 2</td>
<td>0.468</td>
<td>0.094</td>
<td>0.147</td>
<td>0.088</td>
<td>0.497</td>
</tr>
</tbody>
</table>

Note: * p<0.05 ** p<0.01 Standard error in parentheses

Columns (3) and (5) are the test results of the mediator role of digital innovation of retail. The results of column (3) show that digital investment can significantly improve the digital innovation of retail businesses; Column (5) shows the positive impact of retail digital innovation on business performance, and it is significant at the level of 0.05. It can be seen from columns (3) and (5) that retail digital innovation plays a partial mediator role between business digital investment and business performance, which is similar to the conclusion of Qi (2020) who studied the digital transformation of manufacturing industry, and H3 has been verified.

Columns (4) and (5) are the test results of the mediator role of business innovation outside retail. Among them, column (4) indicates that the digital investment of retail businesses can significantly improve the business innovation outside retail, which is also consistent with the conclusion of Qian (2021); In column (5), the coefficient of business innovation outside retail is positive at the level of 0.05. According to columns (4) and (5), it can be seen that the business innovation outside retail has played a partial mediator role, that is, the digital investment of retail businesses has improved the business innovation outside retail and thus improved the performance of businesses, so H4 has been verified.

5 ROBUSTNESS TEST

In the research of management science, the sample itself is highly heterogeneous, and many variables can’t be observed and measured accurately, which is often plagued by endogenous problems. In this paper, there is also an endogenous influence between digital investment and business performance. On the one hand, the degree of digital investment will promote the
performance of businesses; on the other hand, the performance of good businesses will also increase the degree of digital investment. In order to solve the endogenous problem, using the practices of scholars in Wang (2017) for reference, and using lagging explanatory variables for regression, this paper adopts a lagging digital investment.

Table 6: Robustness test.

<table>
<thead>
<tr>
<th></th>
<th>Business performance</th>
<th>Earnings per share</th>
<th>ROE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digitization investment</td>
<td>3.614*</td>
<td>(1.747)</td>
<td>0.809**</td>
</tr>
<tr>
<td>Lag one phase of digital investment</td>
<td>0.059**</td>
<td>-0.537</td>
<td>0.047</td>
</tr>
<tr>
<td>Constant</td>
<td>(0.014)</td>
<td>(0.311)</td>
<td>(0.044)</td>
</tr>
<tr>
<td>Company age</td>
<td>0.002**</td>
<td>0.032**</td>
<td>0.004**</td>
</tr>
<tr>
<td>Cash ratio</td>
<td>0.006</td>
<td>0.154</td>
<td>-0.001</td>
</tr>
<tr>
<td>Gearing ratio</td>
<td>-0.143**</td>
<td>-0.095</td>
<td>-0.182*</td>
</tr>
<tr>
<td>Total assets</td>
<td>0.002**</td>
<td>0.043</td>
<td>0.004**</td>
</tr>
<tr>
<td>Year fixed effect</td>
<td>control</td>
<td>control</td>
<td>control</td>
</tr>
<tr>
<td>N</td>
<td>247</td>
<td>287</td>
<td>287</td>
</tr>
<tr>
<td>R²</td>
<td>0.453</td>
<td>0.236</td>
<td>0.301</td>
</tr>
</tbody>
</table>

Results As shown in Table 6 (1), the regression coefficient is still significant at the level of 0.05, indicating that the significant positive impact of digital investment on business performance still exists after controlling endogenous problems. In addition, in order to ensure the robustness of the model, this paper also adopts new variables to replace the performance level of businesses, which are measured by earnings per share and ROE. The empirical results are shown in columns (2) and (3) in Table 6. The digital investment of businesses has significantly improved the level of EPS and ROE of businesses (the regression coefficients are 3.614 and 0.809 respectively, which are significant at the level of 0.05 and 0.01 respectively). Therefore, the conclusion that digital investment has a positive impact on the performance of retail businesses is robust.

6 CONCLUSION

Based on the data of A-share listed retail businesses from 2010 to 2020, this paper empirically verifies the mechanism of the impact of digital investment on the performance of retail businesses under the background of digital transformation. This paper innovatively measures the digital operation and management of retail businesses. According to the fact that the digital transformation of retail businesses is the transformation of circulation and supply system driven by digital technology with consumer demand as the core, three paths of digital investment's
impact on retail businesses' performance are summarized, and the following conclusions are obtained: First, digital investment has a significant positive impact on retail businesses' performance, and the conclusion is still valid even after one period lag. Secondly, the digital investment has effectively improved the digital operation and management, digital innovation of retail and business innovation outside retail of retail businesses. Thirdly, digital operation and management, digital innovation of retail and business innovation outside retail have significant mediator effects between digital investment and retail business performance, revealing the ways of digital investment's influence on retail business: digital investment→digital operation and management→business performance improvement, digital investment→digital innovation of retail→business performance improvement, digital investment→business innovation outside retail→business performance improvement.

Based on the analysis of the above conclusions, the following implications are obtained:

Increase digital investment and drive business digital transformation with digital technology. Digital transformation is the general trend under the new round of scientific and technological transformation and industrial transformation. Digital technology and digital elements have become the core engines of digital transformation. In 2020, the "14th Five-Year Digital Economy Development Plan" issued by the State Council emphasized the importance of digital technology to promote industrial digitalization. The empirical results also show that digital investment has a positive effect on business performance. Therefore, businesses should follow the trend of development, actively embrace the cutting-edge digital technology, integrate digital technology with business production management activities and business scenarios, and realize the digital upgrade of business organizational structure, business processes, management decisions and business models.

Enhance the application of digital technology to business management and supply chain management, and improve the efficiency of business management and supply chain operation. The empirical results show that digital investment improves the digital operation and management of retail businesses, thus improving business performance. Therefore, retail businesses should make targeted digital investment, use digital technology to promote internal information exchange and knowledge sharing, promote digital improvement of business processes, reduce operating costs, improve business efficiency and gain digital competitive advantage. At the same time, we will invest in Internet of Things, cloud computing, artificial intelligence, blockchain and big data analysis technologies to build an efficient, flexible, transparent and intellectualized supply chain, and realize sales forecasting, intelligent replenishment, logistics route layout and intelligent analysis and decision-making to cope with the rapidly changing market demand, thus reducing warehousing and logistics costs and increasing business sales.

REFERENCES

Big Data Used for Accurately Subsidize Research and Practice in University

Bin Xu

E-mail: 1090573306@qq.com

Wuhan University of Science and Technology, Wuhan, Hubei province, China

Abstract: Targeted subsidize in colleges and universities is a concrete practice of the thought of "targeted poverty alleviation" in the field of colleges and universities, which provides a new concept and path for improving the precision of poverty alleviation in colleges and universities. Big data has the advantages of information collection and analysis, which can provide sufficient information for funders, provide a factual basis for the diversification of subsidize forms, and provide technical conditions for the improvement of subsidize efficiency. In order to realize the goal of optimizing the precise subsidize mode of colleges and universities, we should focus on the integrity of the system, under the guidance of the system theory thought, pay attention to the internal coordination, and construct the theoretical model and institutional innovation from four aspects: framework design, institutional guarantee, technical path and linkage mechanism.

Keywords: Big Data, Precise Subsidize from Universities.

1 INTRODUCTION

General Secretary Xi Jinping mentioned that we should adhere to the targeted poverty alleviation strategy and eliminate the root causes of poverty through development methods. "In poverty alleviation, precision, precision" and "appropriate medicine, precise drip irrigation and targeted treatment". The essence of subsidize is to educate people. subsidize work is related to the development of every independent individual and the fairness and justice of education. It is more necessary to implement accurate thinking. In order to achieve "accurate identification, accurate customization, accurate drip irrigation, accurate evaluation" as the goal, by means of information, data, standardize family economic difficulties students identified procedures, strengthen the subsidize system construction, perfect subsidize education platform, implement fine management, build a set of effective use of big data technology to realize accurate subsidize work mode innovation and carrier construction, effectively improve the subsidize work efficiency, more fully embodies the value of subsidize education.

1.1 Problems Existing in College Student Financial Aid Work at the Present Stage

Lack of timeliness of the subsidize work. In the college student financial aid work, the information of the financial aid object is usually static, and the universities can not obtain the dynamic information in real time, so the student financial aid work often lacks the timeliness.
At present, the application form for the identification of family economic difficulties provided by the students and the certificate issued by the civil affairs department are used by the students as the main basis for student financial aid in colleges and universities. Part of the proof materials can be reused, submitted once, valid for four years. In short, the college subsidize work cannot realize the timely and accurate identification of students from poor families.

Lack of fairness in the subsidize process. The process of college student financial aid is easy to be disturbed by human subjective factors and lacks fairness and objectivity. At present, there is no national poverty alleviation database in China, and schools cannot have an accurate understanding of students' family conditions. In the process of identifying students from poor families and the preliminary evaluation of financial aid, colleges and universities usually adopt the methods of student independent declaration, counselors and managers, and students 'democratic supervision and feedback, and lack of quantitative evaluation of students' real family living conditions and income and expenditure. This way of evaluation is highly subjective, and it is easy to be unfair in the level identification of students from poor families, leading to the phenomenon of "should help but not help" and "not poor but help", and it is difficult to make the subsidize funds play the best allocation benefit [1].

Lack of pertinence of subsidize methods. The subsidize method of colleges and universities is not rich enough and not targeted. At present, the methods of subsidize are mostly grants, scholarships, student loans, green channel setting and work-study positions setting. Due to the lack of data on students' needs, both freshman and senior students implement a unified subsidize method, and there is no targeted assistance to according to the actual needs of different stages and different types of students [2].

Lack of effectiveness in subsidize and education results. In the financial aid, colleges and universities usually support students from the material economy, ignoring the humanistic care for students from poor families, emphasizing poverty alleviation rather than support ambition, there is a phenomenon of lack of financial aid and education effect. Some students with financial difficulties will also have psychological, interpersonal communication and other difficulties, they lack of self-confidence, will often self-denial. Due to the lack of ideological value guidance to the financial aid recipients, some students from poor families lack the sense of gratitude, and think that it is granted to accept the financial support from the state and the society, and lack the sense of return to the society.

1.2 The Role of Big Data for Precise Subsidize from Universities

As a new technology and new method, big data can mine the most valuable data in massive data. By integrating, analyzing, interpretation, decision-making, tracking and prediction of the dynamic data of the sponsor, we can accurately grasp the thought and behavior data of the sponsor and obtain the information we need. The basic information, family data, portrait data, educational administration data, consumption data, grant data and attendance data of students mastered by universities can fully help the realization of big data technology. The use of big data to drive the subsidize work of universities can realize the transformation and upgrading from traditional subsidize to precision subsidize.

Accurate identification and promote accurate identification. Accurate identification of subsidize objects is the premise and foundation of realizing university subsidize. Through the
effective integration of data, big data can maximize the scientific nature of the identification way. First, Big data can be used to master the students’ actual family situation, Test whether the relevant certification materials submitted by the students are true; next, In the identification process, Use big data to master students' basic consumption in study and life, Ensuring the objectivity of the certification process, To overcome the subjectivity of democratic evaluation and school recognition, To achieve the "two one" goal of "no one should be subsidize for students or less" and "no one can" that should not be subsidize for students; Finally, on the level of recognition, By making accurate identification of students' basic information, Knowing the general situation of the subsidize recipients, And to analyze the special circumstances of the subsidize recipients, Make quantitative indicators for the physical health, living conditions and family income of students' family members, Develop a high-precision identification and grade system [3].

Precise customization to realize the personalized subsidize method. Through the dynamic and static analysis of students' information resources by big data, universities can timely find out the relevant relations and existing problems of each work in the accurate subsidize work, so as to formulate the subsidize programs in line with groups and individuals, and achieve accurate subsidize. Using big data, personalized training programs can be formulated for students from poor families with different life pursuits, and students can be targeted to participate in different subsidize projects.

Precise drip irrigation to realize the dynamic subsidize process. To achieve targeted poverty alleviation, we need to change the traditional management mode, start from individual students, realize multi-party cooperation, precision and efficiency, and expand the connotation. We will combine economic poverty alleviation with encouraging success, and combine "wisdom" with "ambition". The application of big data technology can realize the effective integration of information technology, making the subsidize work across time and space limitations. Universities can outline the quality of students' comprehensive development according to the law of subsidize work; then cluster the subsidize factor data (including students' card data, achievement data, student data, class attendance data, online behavior data) to form students' "accurate portrait"; again, embedded intelligent computing model to form "talent navigation", promote the comprehensive development "theme" personalized development; finally, through the big data analysis of subsidize life dynamic capture, timely targeted subsidize, to realize the effective supply of subsidize resources and education process of intelligent interactive.

2 MATERIALS AND METHODS

2.1 Using Big Data to Establish a System Design of University Precision Subsidize

Precision subsidize is a new working mode of subsidize composed of multi-elements and multi-systems, mainly including data collection and analysis system, scientific research and judgment and decision-making system, dynamic evaluation and real-time feedback system, management coordination and system guarantee system.

Precise subsidize breaks the limitations and segmentation of the original data of a single department, forms a cross-departmental and cross-departmental big data sharing cloud
platform, collects the whole-chain data, integrates cross-sectional and correlated data, and realizes data extension and mutual integration and communication. On the one hand, connecting the departments of personnel, engineering, youth League committee, educational affairs, finance, general affairs, realize the effective connection between university subsidize management information system and relevant local departments; with the information system of enterprises, social organizations, communities, entrepreneurship parks, forming the full caliber of student data collection and sorting. On the other hand, from the longitudinal time series, students were collected and integrated from before enrollment and after graduation. Student information data scope includes students and family basic situation, poverty level, poverty cause, student one card consumption, student status, class attendance, entrance guard, bank card payments, students' personal integrity file information to better realize the accurate distribution of funds and found that "recessive poverty" and suspected "virtual false identification of " students. The operation mechanism is shown in Figure 1.

![Figure 1](image)

**Figure 1** The Internal Mechanism of the Construction of University Precision Subsidize System from the Perspective of Big Data

Scientific research and judgment and decision-making system. Precise subsidize requires the construction of a scientific judgment and decision system on the basis of the analysis model. The research and judgment system should use cutting-edge scientific and technological achievements such as deep learning theory and artificial intelligence technology, integrate education, psychology, management, sociology and other related theories, aim at the actual growth and development of the subsidize objects, conduct scientific research and judgment, and provide support for accurate subsidize workers. From the perspective of practical point of view, the research and judgment system includes the identification of family economic difficulties, family economic situation analysis, academic status early warning, daily life status capture, social circle classification and identification, interest potential discovery, employment intelligent matching recommendation, psychological condition mapping and diagnosis, etc.
The research and judgment system can carry out comprehensive research and personalized research on the study, life and thought and behavior of students from poor families, prompting us to fully understand the law of students' growth and implement personalized financial aid\(^4\).

Dynamic evaluation and real-time feedback system. Colleges and universities should build a real-time evaluation and feedback system based on the financial aid needs of students from different families in different economic stages of development and the accurate identification of poor students caused by major emergencies. On the one hand, build a dynamic evaluation system to distinguish the subsidize standards, achieve diversified and developmental subsidize, to meet the different subsidize needs of students \(^5\). Colleges and universities, for example, can through students' personal consumption information, engel coefficient, consumption location parameters and other indicators timely identify subsidize object, the economic improving students appropriately lower subsidize level, the economic situation significantly better students no longer subsidize, major changes, family economy in poor students are not subject to reporting time limit in the subsidize system, for during the subsidize violations, fraud, students can cancel or stop subsidize. At the same time, a dynamic psychological evaluation system should be established to understand the psychological trends of students, provide psychological care for students from poor families, and respond to students' needs. On the other hand, a real-time feedback system should be built to achieve accurate and scientific feedback.

Management coordination and system guarantee system. The precise subsidize mode is not only the change of the thinking mode, but also requires the establishment of the management coordination and system guarantee mechanism of all departments, the establishment of an on-campus and off-campus subsidize platform, to realize the comprehensive docking of the data between the off-campus platform and the on-campus learning and engineering system, so as to realize the whole process, all-round and standardization of subsidize management. One is to establish field joint office system, strengthen the top-level design, through the campus personnel, engineering, education, finance, logistics, employment and other kinds of system, build education teaching integration management system, for family economic difficulties students provide personalized academic guidance, psychological support, interest training, skills training, scientific research guidance, innovation entrepreneurship support, improve the social competitiveness of subsidize object. Second, the introduction of ideological and political education, psychology, communication, management and other professional talents, pay attention to the development subsidize education mode construction, build a set of multidisciplinary management education team, in a daily subsidize and education work at the same time, targeted to big data professional and technical personnel in the field of theoretical study, quality promotion and ability training. Third, establish an information system connecting universities and enterprises and institutions, social organizations, banks, community street offices, and entrepreneurship parks, realize the linkage of various educational resources, and establish a practice platform for data sharing and collaborative education between schools and social subsidize resources \(^6\).
3 RESULTS & DISCUSSION

3.1 Big Data is Used for the Practice of Accurate Subsidize System Construction

This study plans to provide small subsidize to students through real-time monitoring of students' meals and daily consumption. We illustrate the data analysis of Wuhan University of Science and Technology from September 2021 to June 2022. During the epidemic period, the campus adopts a closed management policy, and most students eat in the campus canteen, so the data can more accurately reflect the actual economic situation of the students.

Analysis of the current students' consumption situation

A total of 10,000 undergraduate students, including 5,000 boys and girls, were selected, and their consumption in the campus canteens and supermarkets was counted. According to statistics, the number of consumption times of boys and girls in campus restaurants and supermarkets are shown in Table 1. The average monthly consumption amount is 825.6 yuan. Figures 2-5 shows the average monthly consumption of male and female students of Wuhan University of Science and Technology in canteens and supermarkets during the statistical period.

<table>
<thead>
<tr>
<th></th>
<th>Breakfast frequency</th>
<th>Lunch frequency</th>
<th>Dinner frequency</th>
<th>supermarket consumption frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>male students/m</td>
<td>12.56</td>
<td>20.43</td>
<td>14.65</td>
<td>26.56</td>
</tr>
<tr>
<td>Female students/m</td>
<td>17.45</td>
<td>23.87</td>
<td>17.46</td>
<td>34.62</td>
</tr>
</tbody>
</table>

Figure 2 monthly consumption amount of male students in the canteen

Figure 3 monthly consumption amount of female students in the canteen
Data analysis

According to the above data analysis, the average consumption of boys in the canteen is more than that of girls, but less than that of girls. Girls spend more and average in the supermarket than boys. This may be because boys generally eat more than girls, who are more willing to spend money in canteens and supermarkets.

According to the survey, students with poor family level will give priority to the canteen dining. Therefore, we believe that the more meals in the restaurant during school, the more it can reflect the overall economic level of students. Considering the closed campus management during the epidemic, more students choose to spend money in campus canteens and supermarkets, so the amount of consumption for each meal. In addition, in order to avoid the consumption of less restaurant and more supermarket consumption, we also need to take the consumption of the supermarket as a reference factor. According to the average number of restaurant consumption and amount of students, we determine the standards for poor students are as follows: boys: 50 times, monthly consumption of supermarket 250 yuan, 8 yuan; girls: 60 times, monthly consumption of supermarket 300 yuan, 7 yuan.

Analysis the results of the precision subsidize program for poor students

According to the above conditions, 854 boys and 1025 girls were selected, a total of 1879. We cross-compared the list of the above personnel with the list of students identified as economic difficulties in 2021. The 1,879 people have been identified as poor students, accounting for 86.48% of the financial aid. This shows that our school students in 2021 students that the overall situation is consistent with students' consumption level, but also shows that in the overall financial situation in the process of the implementation, there are still a part of the
students because of various reasons not included in the school subsidize system, or some students family in the recent encounter sudden economic difficulties. In addition, there are also some students in the identification list of poor students, but their consumption level is far beyond the average level, indicating that the identification work of poor students still needs to be further precise and meticulous.

4 CONCLUSIONS

At present, the work is to track, record and analyze each student's information with the help of big data technology, through the real-time monitoring and processing of students' dining, daily consumption and other data, to judge the students' family economic level and school consumption. In the future at the same time through the students in and out of the dormitory time entrance guard information, borrowing library materials, academic data for big data analysis, the comprehensive evaluation of students, auxiliary aid work, more accurately cover the economic difficulties and excellent students at the same time, to help family economic difficulties students get humanized subsidize in a timely manner. Finally, after long-term data accumulation and technical analysis, we will master the situation of the school students, gradually formed the Wuhan university of science and technology students status database, establish student for each difficult student archives, using scientific and effective way to subsidize object of accurate identification, fine management, accurate subsidize.

REFERENCES

Research on Text Recognition System of Logistics Enterprise Policy Based on Text Mining

Jiahui Wang\textsuperscript{1,a}, Yidi Wang\textsuperscript{1,b}, Limei Xu\textsuperscript{1,c}\textsuperscript{*}

wangjiahui96628@163.com\textsuperscript{a}, 1119076056@qq.com\textsuperscript{c}\textsuperscript{*}

College of E-commerce and Logistics, Beijing Technology and Business University, Beijing, China\textsuperscript{1}

Abstract: Modern logistics is an important support to realize the reform of supply side structure and the high quality development of economy. The Chinese government has also formulated a large number of policies to ensure the good and orderly development of logistics. Logistics enterprises concerned about the government policy can grasp the industry wind direction, good business decisions. However, logistics industry is a complex industry, with a wide range of policies, a large number of policies and complex contents. Therefore, enterprises are prone to omissions or inadequate grasp of key fields in practice. Therefore, this paper designs a policy text recognition system for logistics enterprises based on text mining. TF-IDF algorithm is used to extract the feature words of policy texts, and random forest is used to classify policy texts. The results are compared with manual labeling results to calculate the accuracy and recall rate. Through experiments, it is found that random forest algorithm has a high accuracy rate of policy text recognition. Logistics enterprises can use random forest algorithm to identify and analyze policy text, so as to improve the working efficiency and decision-making accuracy of logistics enterprises.

Keywords: Text Mining, Logistics Policy, Analysis System.

1 INTRODUCTION

The total amount of social logistics increased from 219.2 trillion yuan in 2015 to 335.2 trillion yuan in 2021. Logistics has become a strong growth point of the national economy and an important driving force to promote economic development. Especially in the current situation, it is of great strategic significance to vigorously support logistics development to cope with the challenges of economic development under the epidemic, and to establish a new development pattern with the domestic cycle as the main body and the domestic and international cycles mutually reinforcing. With the rapid development of Internet technology and the growing maturity of new generation of basic technologies such as big data and the Internet of Things, logistics has also merged with other industries to give birth to various new models. However, because of the complex property of the logistics industry, the development of the logistics industry was in the state of disorder for a time, which aroused the government's strong
attention. In 2009, The State Council pointed out in the Logistics Industry Adjustment and Revitalization Plan that the legislative research and policy formulation of the logistics industry need to be strengthened, and the corresponding policy and regulation system needs to be improved and improved.

After more than ten years of development, government departments at all levels continue to introduce various policies to guide and support the logistics industry. Studies have proved that the time required for logistics enterprises to adapt to the policies is getting shorter and shorter. Meanwhile, the implementation of the policies has a positive effect on the stock market, market value of logistics enterprises and the development of the logistics industry [1-2]. However, the current policies are numerous, complex, and involve a wide range of industries. Most departments of logistics enterprises collect and apply the corresponding policies manually, which is inefficient and easy to be neglected or repeated between departments in practice. They cannot keep up with the national policy direction and miss some subsidies and support measures. To some extent, this will also affect the decision-making accuracy and future development of logistics enterprises. Therefore, in view of the characteristics of the current logistics policy update fast, large quantity and complex industry, it is necessary to explore the automatic capture and distribution of policy text, establish real-time identification and coordination mechanism, which is conducive to better decision-making and practice of logistics enterprises.

In the era of big data, text data is also growing exponentially. How to capture valuable information in numerous text data is also the research direction of many scholars. Feldman put forward the concept of text mining in 1995 [3]. Text mining mainly uses machine learning techniques such as decision tree and deep neural network to extract text information and analyze text by training machines [4]. After a lot of practice, text mining is widely used in social media because of its high accuracy and speed in text feature analysis, cluster analysis, text emotion and topic extraction [5], policy analysis [6] Literature metrology [7] and other fields of study. Tang Heng [8] textrank method was adopted to extract keywords in the text of intellectual property policies for smes and calculate the weight of each keyword, which directly reflects the focus of intellectual property policies for smes. Ding Siyuan [9] According to the content characteristics and discourse characteristics of public hearing texts, et al. put forward a three-stage event extraction method to realize the extraction of valuable information and conduct in-depth analysis of the extracted valuable information.

In summary, guided by the needs of logistics enterprises for policy interpretation, this study intends to propose an identification mechanism of text classification and feature extraction based on text mining to conduct feature classification and automatic analysis of policy texts, so as to improve the attention and correctness of logistics enterprises to policies, and help logistics enterprises to make correct decisions and develop business strategies consistent with the national wind direction.
2 DESIGN OF POLICY TEXT RECOGNITION SYSTEM

The policy text identification system of logistics enterprises is responsible for automatic identification and distribution of policy texts, and value analysis and processing after processing by various departments, as shown in Figure 1.

The policy text recognition system of logistics enterprises is based on text classification, including data crawling and text feature analysis. Due to the complex nature of logistics industry, policy texts come from many sources and involve a wide range of contents. If manual collection and analysis are carried out, the efficiency is low and the needs of different departments are different, resulting in poor communication. Therefore, this system is based on python and adopts crawler program to collect policy text data, which can establish a general logistics policy text database for logistics enterprises. The policy texts of logistics enterprises are mainly from the state and government functional departments at all levels, such as provinces, cities and counties.

```
print('Grabbing data....')
data = list()
for u in urls:
    try:
        res = requests.get(u, headers=headers)
        html = res.text
        soup = BeautifulSoup(soup, html.parser)
        data = {}
        data['url'] = u
        data['title'] = soup.find('h1').text
        data['time'] = soup.find('time', itemprop)
        if soup.find('div', id='article_body') is
```

Figure 1. Policy text recognition system of logistics enterprises

Figure 2. Code used for data crawling (part)
The policy text identification system of logistics enterprises mainly has the following steps:

Firstly, the original policy text data is preprocessed for word segmentation before extraction and analysis. In this paper, the current mainstream Chinese word segmentation tool - jieba segmentation is adopted for word segmentation. In order to ensure the accuracy and comprehensiveness of word segmentation, this paper uses "HIT University of Technology Glossary of Words" to remove words of words, and introduces "Standard Terms of Logistics (GBT18354-2021)" to build a dictionary of special terms for logistics. In feature engineering, the word bag model is used to represent the text in vector form. The word bag model constructs all the entries in the text data set into a dictionary, and represents each text as a frequency set of entries. The crawler program was used to collect the relevant policy text data, delete the invalid information and use the constructed dictionary for word segmentation of the original policy text.

The TF-IDF algorithm is used for keyword analysis of the policy text data after cleaning, and the category differentiation ability of the entry is explained by calculating the word frequency and reverse file frequency, so as to determine whether the entry is the keyword of the text. It tends to filter out common words and retain important words.

Where, TF is word frequency, which represents the frequency of entry appearing in the document; IDF is the reverse file frequency, indicating the frequency of the file in which the term appears in the whole file set. Multiply the two values together to get the TF-IDF value of the word. The greater the TF-IDF value of the word, the more important it is to the article. It can be expressed by the following formula:

\[
TF = \frac{\text{count}(tj)}{\text{count}(dj)} \\
\text{IDF} = \log \frac{N}{\text{num}(t) + 1} \\
\text{TF-IDF} = TF \times IDF
\]

Where, \(\text{count}(tj)\) represents the number of words \(t\) contained in document \(j\); \(\text{count}(dj)\) represents the total number of words contained in document \(j\); \(N\) is the total number of documents; \(\text{num}(t)\) represents the number of documents containing words \(t\).

The TF-IDF value of each word in the processed database is calculated one by one, and the TF-IDF value set form of each word is used to represent each document, and the irrelevant data is removed, and finally a matrix form document set is obtained.

After understanding the relevant policy texts, this paper takes 106 logism-related policy texts issued by the official websites of The State Council and its subordinate departments from January 1, 2018 to December 31, 2021 as data sources, and marks the department keywords according to the characteristics of logistics enterprises, as shown in Table 1.
### Table1 Department - Keywords in the text of logistics policy

<table>
<thead>
<tr>
<th>The serial number</th>
<th>Department</th>
<th>Labeling Keywords</th>
<th>Number of texts</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Administration and Personnel Department</td>
<td>Talent recruitment, vehicle management</td>
<td>14</td>
</tr>
<tr>
<td>2</td>
<td>Finance Department</td>
<td>Tax administration, corporate income tax, value-added tax, subsidies</td>
<td>52</td>
</tr>
<tr>
<td>3</td>
<td>Management Department</td>
<td>Enterprise strategy, logistics park, pilot, key planning, cultivation, industrial chain, operation mode</td>
<td>98</td>
</tr>
<tr>
<td>4</td>
<td>Department of Transportation</td>
<td>Combined transportation, airport, hub layout</td>
<td>76</td>
</tr>
<tr>
<td>5</td>
<td>Warehouse management Department</td>
<td>Inventory, material turnover, agglomeration, cross-border</td>
<td>69</td>
</tr>
<tr>
<td>6</td>
<td>Research and Development Department</td>
<td>Technical support, network platform, standardization, infrastructure, intelligent logistics equipment</td>
<td>43</td>
</tr>
<tr>
<td>7</td>
<td>Marketing Department</td>
<td>E-commerce, rural logistics, application</td>
<td>87</td>
</tr>
</tbody>
</table>

Classification model can map data records in the database to a given category, including decision tree, logistic regression, naive Bayes, neural network and other algorithms. As a newly emerging and highly flexible machine learning algorithm, random forest has the highest accuracy among all the current algorithms. Besides, it is able to evaluate the importance of various features in classification problems and process input samples with high-dimensional features, so it has obvious advantages in estimation and inference mapping \(^{(10)}\). CART classification regression tree is a typical binary decision tree. In this paper, CART tree is used as a weak classifier, and sklearn standardization tool is used to establish and train the classifier.

### 3 EXPERIMENT OF POLICY TEXT IDENTIFICATION OF LOGISTICS ENTERPRISES

80% of the 106 logistics policy texts collected were used as training texts to establish and train the classifier, and 20% were used as test texts to test the effect of the classifier. The main steps include: segmentation of training data and test data, extraction of dictionary feature vector-quantization, and random forest prediction.
In the random forest model, \( \text{mtry} \) and \( \text{ntree} \) are two important parameters. \( \text{mtry} \) is the number of variables contained in each decision tree, which is generally tried one by one to achieve a better value. \( \text{ntree} \) is the number of base classifiers included. Generally, the value when the error in the model is stable can be roughly judged by the graph. Through the experiment, it is found that a better result can be obtained when the \( \text{mtry} \) value is 7. The setting of \( \text{ntree} \) should make the overall error rate of random forest stable, so the value of \( \text{ntree} \) should be large enough to ensure RF convergence. The results show that the classification effect is best when \( \text{mtry}=7 \) and \( \text{ntree}=20 \). The experimental results are shown in Table 2. The text classification model based on random forest algorithm achieves the best classification effect when \( \text{mtry}=7 \) and \( \text{ntree}=20 \).

Manual labeling has been carried out in the previous paper, and the comparison with the results of the test text classification by the classifier shows that the accuracy rate of random forest classification is above 98%. In order to further verify the accuracy of the classifier, this paper also uses the departmental recall rate for evaluation. The results are shown in Table 3.

Departmental recall rate

\[
R_i = \frac{x_i y_i}{\sum_{i=1}^{7} x_i y_i}
\]  

(4)

The value of \( i \) ranges from 1 to 7 for each of the seven departments.

<table>
<thead>
<tr>
<th>( \text{ntree} )</th>
<th>( \text{precision} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.744</td>
</tr>
<tr>
<td>15</td>
<td>0.753</td>
</tr>
<tr>
<td>20</td>
<td>0.761</td>
</tr>
<tr>
<td>25</td>
<td>0.757</td>
</tr>
</tbody>
</table>

**Figure 3** Model training code (part)

```python
import collections
import pickle
from tools import load_data, word_to,
from sklearn.model_selection import
test from sklearn.feature_extraction imp:
test from sklearn.ensemble import RandomForest

all_data = load_data("train.xlsx")

all_content = list()
for content in all_data["content"]: 
    temp = dict()

  Table 2 Experimental results

<table>
<thead>
<tr>
<th>( \text{ntree} )</th>
<th>( \text{precision} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.744</td>
</tr>
<tr>
<td>15</td>
<td>0.753</td>
</tr>
<tr>
<td>20</td>
<td>0.761</td>
</tr>
<tr>
<td>25</td>
<td>0.757</td>
</tr>
</tbody>
</table>
```
It can be seen from Table 2 that the accuracy of random forest is relatively high, and the classification accuracy of financial department policy text even reaches 100%. It can be considered that random forest can be practiced in the policy text recognition system of logistics enterprises.

4 CONCLUSION

Based on the complex characteristics of logistics industry and from the perspective of logistics enterprises, this paper uses text mining technology to put forward the design idea of policy text recognition system of logistics enterprises, and uses random forest to conduct classification experiment. The experimental results show that this algorithm is effective for the recognition of policy text and department matching of logistics enterprises. The system has a certain practical value, logistics enterprises timely grasp the relevant policies, adjust the company's operation direction, make the right strategic decision plays an auxiliary role, help logistics enterprises to develop in high quality. In the future research, the keywords can be subdivided and improved according to the needs of various departments of logistics enterprises, and more data sets can be introduced for training. Compared with other algorithms, the identification accuracy rate and department matching effect can be better.
REFERENCES


Typical practice of digital management in electrical industry: Characteristics and Cases of International Energy Internet Practice

Jing Wang\textsuperscript{1a}, Yuwei Wang\textsuperscript{1b}, Jiachen Wang\textsuperscript{2c} and Guanghui Wu\textsuperscript{3d}

\textsuperscript{a} 13904396699@139.com, \textsuperscript{b} yuweiwncepu@163.com, \textsuperscript{c} 15851374321@139.com, \textsuperscript{d} 13951603230@139.com

\textsuperscript{1}Economic & Technical Research Institute of State Grid Jilin Electric Power Co., Ltd. Changchun, Jilin, China

\textsuperscript{2}Northeast Dianli University, Jilin, Jilin, China

\textsuperscript{3}Jiangsu Liangjing Technology Co., Ltd. Nanjing, Jiangsu, China

Abstract: Energy Internet is a typical practice of digital management in electrical industry. Energy Internet has formed successful experiences in developed countries, such as Germany, Italy, Denmark, the United States, and Japan. The development process of international energy Internet has been sorted out in this paper is to make reference for the development of energy Internet in other countries. The energy Internet practice exploration of typical countries has been summarized, which is based on their national conditions and their energy system characteristics. The energy Internet demonstration projects in various countries and the energy Internet business carried out by typical enterprises have been studied in this paper.

Keywords: Digital Management, International Energy Internet, Development Practice.

1 INTRODUCTION

Energy Internet is centred on the power system, based on the smart grid, and linked by the Internet, big data, cloud computing, and other latest communication technologies [Tian 2015]. As a new generation of the energy system that comprehensively utilizes power electronics technology and intelligent management technology, Energy Internet can achieve a high degree of integration of energy and information with horizontal multi-source complementarity and vertical coordination of optimization of source, network, load and storage.

Energy Internet is an advanced stage of energy system development, and it is particularly critical and urgent to promote the construction of a clean, low-carbon, safe and efficient energy system. At present, the energy Internet has formed practical cases and successful experiences in Europe, the United States, and Japan. Chinese energy companies such as State Grid, China Southern Power Grid, and China Huadian Corporation LTD. have also started practical explorations in some areas, and some pilot projects have been put into operation [Dong 2014, Wang 2018].
1.1 Characteristics of Energy Internet Development Practice

The development of the international Energy Internet is in its initial stage. The exploration of each country focuses on its national conditions and energy system characteristics and has not formed a unified Energy Internet development model [Wu 2016, Cao 2014].

In terms of development direction, different countries have a highly unified understanding of the basic characteristics of energy Internet, such as multi-energy coupling, flat and decentralized, and high-proportion renewable energy access, such as resilient energy Internet in the United States, the industry coupling energy Internet in Germany, the hydrogen energy Internet in Japan, and the regional energy Internet based on heating technology in Denmark [Yu 2016].

In terms of technological progress, the innovative application of new technologies and elements is valued, including advanced regional cogeneration, hydrogen fuel cells, the Internet of Things, energy blockchain, virtual power plants, etc. Among them, regional cogeneration technology and hydrogen fuel cell technology have achieved good application effects.

In terms of development practice, energy Internet demonstration projects generally focus on the user side, focusing on advanced concepts, key technologies, and elements of regional energy Internet, all of which are combined with national conditions, concepts, and technologies, such as EUREF Campus, EnergyLab Nordhavn project in Denmark, Harumi Olympic Village in Japan [Feng 2017, Wang 2017, Zhao 2020].

1.2 Contents of this paper

The development process of international energy Internet has been sorted out in this paper. The energy Internet practice exploration of typical countries has been summarized, which is based on their national conditions and their energy system characteristics, as shown in Table 1. The energy Internet demonstration projects in various countries and the energy Internet business carried out by typical enterprises have been studied in this paper.

<table>
<thead>
<tr>
<th>Countries</th>
<th>Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Germany</td>
<td>EUREF Campus</td>
</tr>
<tr>
<td>Denmark</td>
<td>The EnergyLab Project</td>
</tr>
<tr>
<td>Italy</td>
<td>Enel</td>
</tr>
<tr>
<td>the United States</td>
<td>Auto Grid</td>
</tr>
<tr>
<td>Japan</td>
<td>Kansai Electric Power Co., Ltd.</td>
</tr>
</tbody>
</table>

2 Development of Energy Internet in Europe

2.1 Overview of Energy Internet in Europe

Developed countries have started building the energy internet at an early stage, with Europe being the first to propose and implement it. It has invested heavily and developed rapidly. Under the EU’s fifth, sixth, and seventh frameworks, European countries have conducted research work on energy co-optimization, integrated energy systems, and other related aspects.
At the same time, they have also carried out a large number of in-depth studies on comprehensive energy development based on their own needs. The development of the European energy internet requires the participation and cooperation of governments, enterprises, academic institutions, and the public at large. Currently, the EU has proposed a series of policies and action plans to promote the development of the European energy internet, such as the “2030 Climate and Energy Framework” and the “European Clean Energy Plan.” Many European countries are also actively exploring and practicing relevant technologies and policies, such as Germany’s “E-Energy” project and France’s “Econetworks” project.

The UK government has been promoting the development of the energy internet by implementing policies and plans to encourage innovation and growth in this field. For example, the UK government launched the Energy Transformation Plan in 2019, which aims to increase the share of renewable energy in the country's electricity mix to 40% by 2050. In addition, the UK government has launched several programs to support clean energy development, such as carbon reduction plans and renewable energy funds. The UK is also making significant technological advancements in the energy internet. For example, National Grid, the UK's largest utility company, is researching and developing smart grid technology to achieve more efficient energy distribution and management. Additionally, UK-based start-ups are developing new energy internet technologies, such as blockchain-based energy trading platforms.

From the perspective of pilot projects promoted in EU member states, Germany is the most representative. Germany's practice mainly relies on two "four-year plans", namely the "Future Energy System" project from 2008 to 2012 and the "Smart Energy-Energy Digital Transformation" demonstration project (SINTEG) from 2016 to 2020. Among them, "Future Energy System" E-Energy is a landmark project. After the implementation of E-Energy project, the German government has also promoted projects such as IRENE, Peer Energy Cloud, ZESMIT, and Future Energy Grid.

Germany is a major energy consumer with relatively scarce oil and gas resources. Since the 1990s, it has made reducing dependence on imported energy a key task of its energy strategy, actively promoting energy demand reduction and increasing energy efficiency while developing renewable energy sources to achieve carbon-free energy supply. Germany is one of the earliest countries to explore and practice the energy internet, focusing on digital transformation of the entire energy system to promote the development and utilization of renewable energy sources, thereby promoting energy structure transformation and energy efficiency improvement.

Under the guidance of government departments, Germany has conducted two series of pilot projects in stages, initially exploring and testing the potential and implementation methods of information communication technology for coupling, interconnecting, and trading different energy systems. The E-Energy project launched in 2008 has comprehensively explored the market, technology, and system levels of information and communication technologies that drive the coupling, interconnecting, and trading of different energy systems. The C-sells project launched in 2012 is part of the SINTEG plan and aims to create small digital energy interconnection systems to achieve regional micro-balance and optimization of energy production and sales.
2.2 Cases of Energy Internet in Europe: EUREF Campus, EnergyLab Nordhavn and Enel

EUREF Campus, located in the southwest of Berlin, Germany, is a pilot dedicated to implementing advanced concepts of energy transformation. 80%-95% of the energy needs of the entire pilot are met by renewable energy sources such as wind power, photovoltaic, geothermal, and biogas. Through technologies such as smart microgrid systems, ultra-low energy buildings, and IoT-based control systems, the efficient use of renewable energy is realized, intelligent charging and discharging of electric vehicles, flexible conversion of hot and cold energy storage, and convenient energy transactions.

The EnergyLab Project in Nordhavn, Denmark, is deeply participated by ABB and the Danish University of Science and Technology, which focuses on the practice of several key technologies and concepts, including regional low-temperature heating technology to achieve deep interconnection of electric heating networks, energy service-oriented flexible management platform, Intelligent control of buildings for flexibility. This project has installed intelligent collection and control equipment in 19 apartments in the area. Through the intelligent control of the electrical equipment (mainly air conditioners and heat pumps), it has realized the improvement of system adjustment benefits, the reduction of user energy costs, and energy conservation and emission reduction. In terms of system conditioning benefits, these homes reduced morning peak loads by an average of 68%. In terms of energy-saving benefits for users, the average heating cost is reduced by 15%, and the energy-saving benefits are significant. In terms of system economy and environmental protection, it saves 600,000 DKK and reduces CO2 emissions by about 70t per year.

Enel in Italy covers power generation, transmission and distribution, power trading, gas transmission, gas storage, gas sales, and smart energy services. Its business sectors include global infrastructure and power networks, global trade, and global power production. Affected by the electricity marketization reform carried out in Italy since 1999, Enel divested its power transmission business, focused on its main business, and looked for new business growth points. To meet the Kyoto Protocol and EU 2050 Low-Carbon Economy Roadmap, EGP was established in 2008 to carry out the acquisition, development and management of global renewable energy. In 2009, Enel took precautions to build an urban intelligent lighting system; under the environment of continuous strict control, it increased the development and investment in emerging integrated energy services such as electric vehicles, energy storage, and smart cities. In 2016, Enel cooperated with AWS to realize all the applications online. In 2017, EneX was established to operate independently, providing digital value-added services for industry and commerce, transportation, cities and families. It acquired several companies including EnerNoc and quickly formed a healthy development trend. In 2020, build a digital middle platform and realize the output of standardized products and services related to the user side.
3 Development of Energy Internet in the United States

3.1 Overview of Energy Internet in The United States

The US Department of Energy (DOE) is the highest authority for various types of energy resources in the US and is responsible for developing related energy policies. The US Energy Regulatory Commission is responsible for implementing government energy policies and suppressing the random fluctuations of energy prices. Under this management mechanism, various energy systems in the US have achieved better coordination and cooperation. Meanwhile, typical integrated energy suppliers such as PG&E and Edison Electric have developed well.

The US government has been promoting the development of energy internet through policy and plan formulation to promote innovation and development in this field. In 2015, US President Barack Obama signed the Clean Energy Plan to promote the development and utilization of clean energy, including strengthening the construction of smart grids, promoting distributed energy, and other measures. In addition, the US government has launched a series of plans to support the development of clean energy, such as renewable energy funds, carbon reduction plans, etc. The technology of US energy internet is also constantly improving. For example, the US has started to build decentralized energy trading platforms based on blockchain technology to achieve transparent and secure energy trading and settlement. Tesla is researching and developing advanced battery technology and energy storage solutions to support the development of the electric vehicle market. In addition, some companies are also developing intelligent meters, smart homes, and other products to improve the efficiency and convenience of energy use.

The US already has some practical cases that demonstrate the feasibility and advantages of energy internet. For example, California has established a smart grid project called “CaliSmart” that achieves real-time monitoring and control of the power system, improves the reliability and efficiency of power supply. In addition, states such as Florida and Texas are actively exploring the application of distributed energy and microgrid technologies.

3.2 Cases of Energy Internet in The United States: Auto Grid

Auto Grid in the United States mainly involves big data of energy. The business content includes energy consumption prediction, optimized operation, and demand management. Its development strategy is to create a brain for smart meters. Demand side products include electricity consumption forecast, combined with electricity price to realize demand response and demand side management analysis; smart home products mainly include demand response management, power generation and grid dynamic load prediction and grid operation fault prediction. Auto Grid has established an energy data platform architecture online, and has specially established a public facility backend. It collects and analyzes the data on the energy demand side (industry, enterprises, and residents) through the Demand Response Optimization and Management System (DROMS), and these data were uploaded to the Energy Data Platform (EDP) for analysis, prediction and storage online. The business model includes a SaaS model, shared revenue model and cooperation model. Among them, the SaaS model is that users pay according to the amount of data processed by Auto Grid; the shared revenue model refers to that Auto Grid sends reports to customers, the customer responds to demand,
and shares revenue with customers; the cooperative model refers to providing software to equipment manufacturers, and charge the equipment manufacturer a license fee.

4 Development of Energy Internet in Japan

4.1 Overview of Energy Internet in Japan

Japan is a highly developed industrialized country and also a major energy consumer. With the continuous growth in energy demand and increasingly serious environmental issues, the Japanese government has started to pay attention to the development of energy internet and has taken a series of measures to promote its development. The Japanese government has promoted the innovation and development of energy internet through the formulation of relevant policies and plans. In 2017, the Japanese government released "Smart Society Strategy" and "Green Future Strategy", proposing measures such as promoting smart grids and renewable energy. Japan also has certain accumulation and advantages in energy internet technology. Japan has built a series of infrastructure, such as intelligent meters and electric car charging stations, and actively develops advanced battery technology and energy storage solutions. Many Japanese companies have also started to invest more in the field of energy internet. For example, Tokyo Electric Power Company is developing a decentralized energy trading platform based on blockchain technology, and Toyota Motor Corporation is also researching and developing new energy vehicle technology. Japan also actively participates in international cooperation, working with other countries to promote the development of energy internet. Japan has signed multiple cooperation agreements with countries such as the United States and Europe to jointly promote research and development in areas such as smart grids and electric vehicles.

4.2 Cases of Energy Internet in Japan: Kansai Electric Power Co., Ltd.

Kansai Electric Power Co., Ltd. of Japan covers power, gas, communication and livelihood business service, and its customer groups include large industrial and commercial users, small commercial and household users, etc. Before 2015, Kansai Electric Power mainly focused on the power industry, and the other three sectors (IT, integrated energy supply and livelihood services) were relatively independent business sectors. With the gradual opening up of the electricity retail market, the electricity and gas sales business has been impacted. In 2016, K-Opticom, a communications company under the Group, sold a bundled package of communication and electricity services. In 2018, the Group launched the electric package Nattoku Pack, an electric co-sale package. The Group began to integrate its business, and initially formed a business model with the integrated energy supply of electricity and gas as the core, integrating with the other three sectors and growing together. Driven by the 5D phenomenon of the big social environment, the Group proposed an overall strategy driven by digital transformation, and established a new joint venture company k4 digital with Accenture in due course to promote the transformation.
5 Conclusion

To sum up, the development of the energy Internet in Europe, the United States and Japan mainly adopt top-down promotion. The summary in this paper will provide targeted reference for the development of energy Internet for other countries. In the process of promotion, industry guidance and policy support at the national level are indispensable, focusing on the practical exploration of highlighting key links, core concepts and key technologies, and based on extensive use of Internet technology, they have conducted in-depth cooperation with professional energy data analysis companies.

References

China-ASEAN International Logistics Development Strategy

Tian Fang*

*Corresponding author: fangtianrioh@163.com
Research Institute of Highway Ministry of Transport, Beijing, 101088, China

Abstract: With the increasingly complex international situation, ASEAN is becoming more and more important to China in international trade status, and logistics development is the basis for the development of international trade. Based on the development of ASEAN's economy, the basic situation of China-ASEAN trade and the basic situation of ASEAN's international logistics development, this paper sets up an impact factor model of China-ASEAN international logistics network, combines the logistics service quality level of ASEAN countries, adopts a combination of qualitative analysis and quantitative analysis, and puts forward policy recommendations for China-ASEAN international logistics network.

Keywords: China-ASEAN, International Logistics, Demand Forecast, Comprehensive Transportation.

1 INTRODUCTION

ASEAN is the abbreviation of the Association of Southeast Asian Nations (ASEAN). Since its establishment in Bangkok in 1967, ASEAN has ten member countries, including Indonesia, Malaysia, the Philippines, Singapore, Thailand, Brunei, Vietnam, Myanmar, Laos and Cambodia, with a total area of about 4.52 million square kilometers and a population of 660 million, accounting for 8.6% of the world's total population. As a close neighbor and an important partner of China, ASEAN's trade and investment have grown against the backdrop of the downturn in the world economy. For the second consecutive year, ASEAN has become China's largest trade partner, with closer economic ties and closer industrial ties between the two sides. Smooth logistics is an important guarantee for trade development, so it is of great research value to study and promote the establishment of China-ASEAN international logistics development strategy.

The research content of logistics development strategy at home and abroad is relatively rich, and the principal component analysis method used in this paper is also one of the main research methods. However, in the current logistics research for ASEAN, some studies select major cities as the perspective to study urban logistics nodes\(^1\), but the author believes that the strategic suggestions for logistics development between China and ASEAN should be considered as a whole country. At the same time, in the research of logistics development strategy, scholars
mostly adopt quantitative research methods, but the combination of ASEAN development strategy and the current situation of China and ASEAN international logistics development is insufficient. Therefore, this paper proposes policy recommendations to promote the development of international logistics between China and ASEAN by combining quantitative research with qualitative research, combining macro policy requirements and micro logistics development.

2 FACTORS AFFECTING CHINA ASEAN INTERNATIONAL LOGISTICS

2.1 Economic Development of ASEAN Countries

According to the statistics of the World Bank, the ASEAN economy has shown a steady growth trend in recent years. The total economic volume of ASEAN has increased from 260 million dollars in 2016 to 340 million dollars in 2021, accounting for about 3.5% of the total world economy. Although affected by adverse factors such as the COVID-19, the manufacturing purchasing managers' index (PMI) of ASEAN member countries continues to be in the expansion range. ASEAN countries are highly dependent on foreign trade and the volume of foreign trade in goods accounts for more than half of GDP in most countries.

The economic strength of the ten ASEAN countries is uneven, with a huge gap in GDP and per capita GDP, as well as the import and export trade volume of each country. On the whole, except Malaysia, the economic development of countries in Indochina Peninsula is relatively backward, and the old ASEAN member countries including Indonesia, Thailand, the Philippines, Singapore and Malaysia rank in the top five. In recent years, the total GDP of these five countries accounts for more than 80% of the total GDP of ASEAN.

![The GDP per capita and growth of ASEAN member countries in 2020](image)

*Figure 1.* The GDP per capita and growth of ASEAN member countries in 2020
The industrial structure of ASEAN member countries is quite different. According to the data of the World Bank, in 2020, Singapore's service industry accounted for 71% of GDP, exceeding the world average (about 65%), and Thailand and Malaysia's service industry accounted for nearly or more than 55% of GDP. The proportion of industrial added value in GDP is higher than the world average (about 25%) in all 9 member countries except Singapore. Brunei's industrial added value accounts for more than 59% of GDP. Except Singapore and Brunei, the proportion of agricultural added value in GDP of other 8 member countries is significantly higher than the world average (about 3.6%), and the proportion of agricultural added value in GDP of Cambodia and Myanmar is nearly 23%.\[^2\]

![Figure 2. The industry structure of ASEAN member countries in 2020](image)

### 2.2 Basic Situation of China ASEAN Trade

After China put forward the "the Belt and Road" initiative, ASEAN is the most important fulcrum region and cooperative partner of the "the Belt and Road", and bilateral trade between China and ASEAN also shows a trend of continued rapid growth. In 2021, the volume of goods trade between China and ASEAN will reach 878.2 billion US dollars, with a year-on-year growth of 28.1%. Vietnam, Malaysia and Thailand are China's top three trading partners in ASEAN. The implementation of China ASEAN Free Trade Agreement and RCEP Agreement has further consolidated the economic relationship between China and ASEAN.

There are significant gaps among ASEAN member countries. In terms of countries, Vietnam, Malaysia and Thailand are China's top three trading partners in ASEAN. In 2020, the bilateral trade in goods between China and Vietnam reached 1922.8 US dollars, with a year-on-year growth of 18.7%, accounting for 28.0% of the total trade in goods between China and ASEAN in the same period. Malaysia followed closely, with a trade volume of 131.16 billion US dollars in 2020, up 5.7% year on year. The trade between Thailand, Singapore, Indonesia and the Philippines has developed steadily, while the trade volume between Myanmar, Cambodia, Laos,
Brunei and China is relatively low. From the perspective of trade trend, China's trade with other ASEAN countries is growing, except for the decline in trade with Laos, Indonesia and Singapore.

2.3 Basic Situation of ASEAN International Logistics Development

The functional positioning of different ASEAN regional networks bears different quantities of goods. At present, the ASEAN region is mainly engaged in the transportation of goods by sea, mainly bulk cargo transportation, which is responsible for about 80% of ASEAN's cargo volume. The second is road transportation. The main road ports of Vietnam, Laos, Cambodia, Thailand and Myanmar on the border are all connected by roads, but most of the roads are low grade, mostly Class III and Class IV roads. The railway transportation is being improved in the ASEAN Peninsula countries, mainly responsible for the transportation of industrial raw materials. Due to high cost and small proportion of transportation volume, air transport mainly transports high value-added products.

ASEAN attaches great importance to the transport connection between internal countries. In the 2025 ASEAN General Plan for Connectivity (2016) and other documents[3], ASEAN has made a strategic layout for strengthening the connectivity between countries in terms of aviation, maritime transport, roads, railways, etc. At the same time, ASEAN countries have improved the transport convenience and unity among countries through the ASEAN Framework Agreement on Facilitation of Freight Transport (AFAFGIT), the ASEAN Framework Agreement on Facilitation of Transport among Countries (AFAFIST) and the ASEAN Framework Agreement on Multimodal Transport (AFAMT).

3 CONSTRUCTION OF CHINA ASEAN INTERNATIONAL LOGISTICS NETWORK

3.1 Analysis on the Influencing Factors of China ASEan International Logistics Network

According to the above analysis, determine the indicators that affect the logistics level of ASEAN countries, mainly including the comprehensive development level of the country, the overall logistics demand of the country, the logistics demand of China and the level of comprehensive logistics facilities. Then, determine the correlation degree of corresponding evaluation indicators, and select 16 factors with high correlation degree as evaluation indicators. The specific evaluation indicators are shown in the following table.

<table>
<thead>
<tr>
<th>Level 1 indicators</th>
<th>Secondary indicators</th>
<th>Third level indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>National comprehensive development</td>
<td>Economic development</td>
<td>GDP (100 million US dollars) (X1)</td>
</tr>
<tr>
<td></td>
<td>Development of tertiary industry</td>
<td>GDP per capita (100 million US dollars) (X2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Agricultural added value (100 million US dollars) (X3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Industrial added value (100 million US dollars) (X4)</td>
</tr>
<tr>
<td>Country</td>
<td>X1</td>
<td>X2</td>
</tr>
<tr>
<td>--------------</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>Vietnam</td>
<td>34</td>
<td>34</td>
</tr>
<tr>
<td>Thailand</td>
<td>50</td>
<td>71</td>
</tr>
<tr>
<td>Myanmar</td>
<td>81</td>
<td>15</td>
</tr>
<tr>
<td>Laos</td>
<td>26</td>
<td>55</td>
</tr>
<tr>
<td>Cambodia</td>
<td>33</td>
<td>85</td>
</tr>
<tr>
<td>Malaysia</td>
<td>36</td>
<td>22</td>
</tr>
<tr>
<td>Philippines</td>
<td>34</td>
<td>00</td>
</tr>
<tr>
<td>Singapore</td>
<td>10</td>
<td>59</td>
</tr>
<tr>
<td>Brunei</td>
<td>12</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 2. China ASEAN international logistics network influencing factors
The characteristic root is obtained according to the correlation coefficient matrix. Based on the principle that the characteristic root is greater than 1, the first three principal components are extracted. At the same time, it can be seen from the table that the cumulative contribution rate of the first three principal components is 90.4%, indicating that 90% of the data information is reflected. Calculate the initial factor load of principal components to find out the indicator information reflected by each principal component.

<table>
<thead>
<tr>
<th>Composition</th>
<th>Characteristic value</th>
<th>Percentage of component variance in total variance (%)</th>
<th>Cumulative percentage of variance of each component in total variance (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8.075</td>
<td>50.47</td>
<td>50.47</td>
</tr>
<tr>
<td>2</td>
<td>4.934</td>
<td>30.84</td>
<td>81.306</td>
</tr>
<tr>
<td>3</td>
<td>1.456</td>
<td>9.10</td>
<td>90.403</td>
</tr>
<tr>
<td>4</td>
<td>0.74</td>
<td>4.63</td>
<td>95.031</td>
</tr>
<tr>
<td>5</td>
<td>0.453</td>
<td>2.83</td>
<td>97.865</td>
</tr>
<tr>
<td>6</td>
<td>0.166</td>
<td>1.04</td>
<td>98.905</td>
</tr>
<tr>
<td>7</td>
<td>0.103</td>
<td>0.64</td>
<td>99.548</td>
</tr>
<tr>
<td>8</td>
<td>0.071</td>
<td>0.44</td>
<td>99.992</td>
</tr>
<tr>
<td>9</td>
<td>0.001</td>
<td>0.01</td>
<td>100</td>
</tr>
<tr>
<td>10</td>
<td>0.00</td>
<td>0.00</td>
<td>100</td>
</tr>
<tr>
<td>11</td>
<td>0.00</td>
<td>0.00</td>
<td>100</td>
</tr>
<tr>
<td>12</td>
<td>0.00</td>
<td>0.00</td>
<td>100</td>
</tr>
<tr>
<td>13</td>
<td>0.00</td>
<td>0.00</td>
<td>100</td>
</tr>
<tr>
<td>14</td>
<td>0.00</td>
<td>0.00</td>
<td>100</td>
</tr>
<tr>
<td>15</td>
<td>0.00</td>
<td>0.00</td>
<td>100</td>
</tr>
<tr>
<td>16</td>
<td>0.00</td>
<td>0.00</td>
<td>100</td>
</tr>
</tbody>
</table>

It can be seen from Table 4 that the first principal component has a high load at X6, X7, X8, X9, X12, X15 and X16. The second principal component mainly reflects the information of X1, X2, X3, X4, X5 and X14, and the third principal component mainly reflects the information of X10 and X11. Therefore, extracting three principal components can basically reflect the information of all components.

<table>
<thead>
<tr>
<th>Indicators</th>
<th>Initial factor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>X1</td>
<td>0.648</td>
</tr>
<tr>
<td>X2</td>
<td>0.572</td>
</tr>
<tr>
<td>X3</td>
<td>0.346</td>
</tr>
<tr>
<td>X4</td>
<td>0.552</td>
</tr>
<tr>
<td>X5</td>
<td>0.649</td>
</tr>
<tr>
<td>X6</td>
<td>0.949</td>
</tr>
<tr>
<td>X7</td>
<td>0.952</td>
</tr>
<tr>
<td>X8</td>
<td>0.834</td>
</tr>
</tbody>
</table>
Utilizing the primary component initial factor load matrix and standardized data, the expressions of the first principal component $F_1$, the second principal component $F_2$ and the third principal component $F_3$ are calculated respectively, and the ratio of the corresponding eigenvalue of each principal component to the sum of the extracted principal component eigenvalue is used as the weight to calculate the principal component comprehensive model, as shown in the following expression.

$$F = 0.558F_1 + 0.341F_2 + 0.101F_3$$

Taking the first principal component as an example, the formula for calculating the principal component is as follows:

$$F_1 = \sqrt{8.075}(0.648XZ_1 + 0.572XZ_2 + 0.346XZ_3 + 0.552XZ_4 + 0.649XZ_5 + 0.949XZ_6 + 0.952XZ_7 + 0.834XZ_8 + 0.777XZ_9 + 0.742XZ_{10} + 0.713XZ_{11} + 0.747XZ_{12} + 0.317XZ_{13} + 0.188XZ_{14} + 0.927XZ_{15} + 0.875XZ_{16})$$

<table>
<thead>
<tr>
<th>countries</th>
<th>F1</th>
<th>F2</th>
<th>F3</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thailand</td>
<td>14.491</td>
<td>9.355</td>
<td>-0.021</td>
<td>11.278</td>
</tr>
<tr>
<td>Myanmar</td>
<td>-20.543</td>
<td>1.102</td>
<td>0.365</td>
<td>-11.055</td>
</tr>
<tr>
<td>Laos</td>
<td>-24.141</td>
<td>-5.205</td>
<td>0.476</td>
<td>-15.204</td>
</tr>
<tr>
<td>Cambodia</td>
<td>-22.878</td>
<td>-4.061</td>
<td>0.286</td>
<td>-14.128</td>
</tr>
<tr>
<td>Malaysia</td>
<td>10.893</td>
<td>1.543</td>
<td>-2.204</td>
<td>6.385</td>
</tr>
<tr>
<td>the Philippines</td>
<td>-6.456</td>
<td>1.680</td>
<td>-0.087</td>
<td>-3.040</td>
</tr>
<tr>
<td>Singapore</td>
<td>43.574</td>
<td>-20.751</td>
<td>1.470</td>
<td>17.395</td>
</tr>
<tr>
<td>Indonesia</td>
<td>17.495</td>
<td>21.264</td>
<td>2.491</td>
<td>17.270</td>
</tr>
<tr>
<td>Brunei</td>
<td>-22.939</td>
<td>-7.372</td>
<td>0.876</td>
<td>-15.232</td>
</tr>
</tbody>
</table>

It can be seen from the scores in the table above that Singapore and Indonesia scored much higher than other ASEAN countries, followed by Thailand, Malaysia and Vietnam. Other ASEAN countries scored lower.

### 3.2 Analysis of Logistics Service Quality in ASEAN Countries

Logistics development has been recognized as one of the core pillars supporting trade development. The National Logistics Performance Index (LPI) jointly proposed by the World Bank and the Turku School of Economics in Finland is an accurate and reliable evaluation system. The LPI index ranges from 1 to 5. The higher score means the better performance.
From the perspective of the influencing factors of China ASEAN logistics network and the logistics service quality of ASEAN countries, the overall level of Singapore, Indonesia, Thailand, Malaysia and Vietnam is high, while the development level of the Philippines, Myanmar, Cambodia, Laos and Brunei is low.

4. STRATEGIC SUGGESTIONS ON CHINA ASEAN INTERNATIONAL LOGISTICS NETWORK

4.1 Overall Scheme

Based on the existing international logistics network infrastructure conditions of China and ASEAN, combined with the ASEAN development strategy, according to the above analysis of the logistics network impact ranking of ASEAN countries and the analysis of the logistics service quality level of each country, the ASEAN countries are divided into two groups. Singapore, Indonesia, Thailand, Malaysia and Vietnam are priority development considerations, and the Philippines, Myanmar, Cambodia, Laos and Brunei as supplementary development considerations.

It is suggested to give full play to comparative transport advantages in every country, so as to realize the international transportation between China and ASEAN in a larger scope and with more efficiency. As the construction of ports and railways is determined by planning, the linking role of international road transport should be fully played in the development of comprehensive transport, and the China ASEAN international logistics network should be built.
4.2 Specific Strategic Recommendations

1) Maritime transport
Among the countries with high comprehensive scores, Singapore, Malaysia and Indonesia are all far away from China and have obvious maritime transport advantages. Therefore, the development of international logistics in the above three countries should focus on the opening of key routes in the ocean. Actively connect with Kuala Lumpur Transport Strategic Plan (ASEAN Transport Strategic Plan 2016-2025) [5]. Currently, Indonesia, Malaysia, the Philippines and Thailand were exploring possible avenues to operationalize the 3 recommended priority routes, namely: Santos-Bitung route, Melaka-Dumai route, and Belawan-Phuket route. It is suggested to strengthen China's investment in infrastructure and other related projects at key ports to strengthen transportation between key ports.

2) International road transport
As China is adjacent to some ASEAN countries and has a relatively short transportation distance with neighboring countries, compared with other transportation modes, international road transportation can realize door-to-door, and has obvious advantages in transportation time and cost. The logistics service network of ASEAN countries adjacent to China shall be unblocked first, and the connectivity with ASEAN countries shall be promoted from near to far. Based on the above analysis, it is suggested to promote the construction of China ASEAN international road transport corridor by stages and steps.

   (1) First, develop China-Vietnam international road transport corridor, and pass through Cambodia and Thailand to open the eastern Kunming/Nanning Hanoi-Danang-Nha Trang-Ho Chi Minh City-Phnom Penh-Bangkok.

   (2) Continue to promote the construction of China-Thailand international road transport corridor, and strengthen the infrastructure construction in Laos, and extend it to Malaysia and Singapore, open the middle line to Mohan-Vientiane- Bangkok-Kuala Lumpur-Singapore international road transport route, and accelerate the construction of a transnational "southward" corridor with ASEAN Indochina Peninsula countries running through the north and south.

   (3) Promote the construction of infrastructure in Myanmar and Cambodia, promote the construction of Ruili-Mandalay corridor, provide technical support for Ho Chi Minh Phnom Penh Bangkok line, and complement the weakness of infrastructure construction.

   For out of gauge goods that need to be transported through the cold chain, such as goods, parts transportation and machinery transportation that cannot be disassembled, road transportation has unique advantages that other transportation methods cannot match, and it also conforms to the characteristics of trade development between China and ASEAN. In the development of China ASEAN international road transport, international road transport facilities should be provided for agricultural products, electronic equipment and other special trade materials between China and ASEAN.

3) Railway transportation
In terms of railway transportation, the Singapore Kunming Rail Link (SKRL) project is an integral part of the Trans Asian Railway, with three planning lines: the east line, the middle line and the west line. It was determined in 2006 and planned to be completed in 2036. Up to now, the China Laos line has been completed, and the Kunming Laos Vientiane section has been
opened, which is assisted by China. For other sections, due to different railway gauge, shortage of funds, and different interests of some countries, there is not much progress in other sections.

5 CONCLUSION

ASEAN is of great strategic significance to China in international trade, and international logistics is the basis of trade. This paper starts from the factors that affect China ASEAN international logistics, uses the principal component analysis method to analyze the different logistics situations in ASEAN countries, and combines the logistics service quality of ASEAN countries to draw strategic recommendations for the development of China ASEAN international logistics network. The main conclusions of this paper are:

(1) ASEAN countries are divided into two levels of linkage development, of which Singapore, Indonesia, Thailand, Malaysia and Vietnam are priority development countries, and other countries are complementary development countries. The purpose is to concentrate resource advantages and take advantage of the role of logistics.

(2) According to the division of countries and ASEAN development strategy, specific development suggestions are put forward for maritime transport, international road transport and railway transport. Most of the top countries are remote countries, which should actively support the development of maritime transport and increase the control of ports. According to the comprehensive ranking of ASEAN countries, put forward policy recommendations to promote the development of the international road. As the railway development is relatively stable, it is recommended to develop the combined transport of highway and railway.

REFERENCES

Coordination of Intermodal Transport in Yangtze River Economic Belt in China Based on the Fourth-Party Logistics

Xiaofang Chen¹, Junli Zheng¹*, Junbo Zhao¹

* Corresponding author: zhjl@sjtu.edu.cn

Shanghai Jiao Tong University, 1954 Huashan Road, Shanghai, 200230, China¹

Abstract: With the deepening of economic integration, intermodal transport has become a central issue worldwide in recent years. China also pays great attention to this issue and promotes the development of intermodal transport in Yangtze River Economic Belt in China. However, the coordination of intermodal transport is a multi-sector, multi-agent and multi-mode process and some problems emerge in the development of intermodal transport in Yangtze River Economic Belt. This paper aims to provide some advice for the improvement of intermodal transport in this region. This study first identifies the barriers to the development of intermodal transport from the aspects of policies, actors, terminals, transshipment services, and information systems. In addition, this study proposes that 4PL can promote standardization, coordination, utilization and innovation of information in the intermodal chain. Furthermore, this paper tries to design the 4PL platform and makes suggestions for the development of the 4PL platform in Yangtze River Economic Belt. Meanwhile, the Chinese government is required to enact some supporting policies to encourage 4PL to play the role of coordinator in the intermodal chain.

Keywords: Integrated Transport, Intermodal Transport, Fourth-Party Logistics, Information Platform.

1. INTRODUCTION

In the new global economy, intermodal transport has become a central issue worldwide[1]. Trade globalization has greatly increased the freight flows among different countries. According to the report from Reports and Data, the market share of global intermodal freight transport is expected to reach 73.38 billion by 2026. Most issues adopt the definition of intermodal transport from The European Conference of Ministers of Transport[2], which defines intermodal freight transport as “the movement of goods in a single freight unit through two or more successive modes of transport, with no handling of the freight during transportation”.

China also channels great effort into the development of intermodal transport. In 1970, China introduced the concept of international container transport[3], from which intermodal transport has become a significant transportation mode in China. In 2014, China launched the Intermodal Transport Demonstration Project to improve the integrated transportation system. In 2021, China proposed to promote the integrated development of various modes of transportation in the 14th Five-Year Plan. By 2022, China has conducted the fourth Intermodal Transport Demonstration...
Project and published Measures for Administration of the Intermodal Transport. However, the volume of intermodal freight transportation in China is still growing slowly. Intermodal transport makes up less than 5% of the total freight volume up to now, while the proportion in Europe and US is generally over 30%. Due to the low price of road transportation, the volume of road transportation in China accounts for about 75% in recent years and keeps increasing nowadays[4].

The Yangtze River Economic Belt is the key development region of intermodal transport in China. The Yangtze River Economic Belt covers eleven provinces, which are usually divided into three parts according to the watershed discretization[5]. The government has formulated a three-year action plan to promote the development of intermodal transport in the region in 2018 and the proportion of the Intermodal Transport Demonstration Projects in Yangtze River Economic Belt is up to 37.1% by 2022. That is because the intermodal transport in Yangtze River Economic Belt is fundamental to the integrated development of the regional economy. Intermodal transport extends the coverage of services of seaports by connecting the seaports with inland ports, which contributes to driving the industry development of inland areas. Additionally, intermodal transport gives full play to the overall advantages of the industrial division of labor and coordination. It is an opportunity to build the cross-region industrial transfer pattern and foster the important growth poles in China through the match of the transportation mode and the industry layout. Furthermore, intermodal transport also leads to three-industry integration as it improves the distribution of agricultural and industrial products. With the reduction of unnecessary transshipment operations, the supply chain has been improved and the costs are reduced.

Figure 1. Policies of intermodal transport in China
2. RELATED LITERATURE

More recent attention has focused on integrated transport. The concept of integrated transport has become an important guiding rule of transport policies in several countries\(^6\). Various studies have tried to clarify the definition and features of integrated transport. NEA7 defines integrated transport as “the organizational process through which the planning and delivery of elements of the transport system are brought together, across modes, sectors, operators and institutions, with the aim of increasing net social benefits”. Eggenberger and Partidário8 propose five forms of transport integration, which are “substantive, methodological, procedural, institutional and political”. In general, integrated transport often relates to the term “intermodal”\(^9\), which is an important part of the integrated transport network. In addition, some issues study the importance of integrated transport. Booth and Richardson10 state that transport planning is a process that involves multiple agencies, sectors and modes, thus the balance of different interests, issues and policies is needed. However, operational managers sometimes focus more on a limited sub-set of the transport system\(^9\).

As a critical part of integrated transport, a large and growing body of literature has studied intermodal transport. Some literature also uses the term “multimodal transport”, but “intermodal transport” and “multimodal transport” are generally used interchangeably\(^2\). Agamez-Arias and Moyano-Fuentes2 identify three research lines of intermodal transport which include basic principle, way of improvement and system modeling. Nowadays, most literature pays attention to the optimization of the intermodal transport system as various decision problems demand the operation research techniques\(^11\). Zhang et al.\(^12\) establish a collaborative planning model considering the collaborations among carriers and propose an Adaptive Large Neighborhood Search heuristic method\(^12\). Abbassi et al.\(^13\) suggest a robust optimization model for dealing with uncertainties of using costs, transport costs and capacities of terminals\(^13\).

The concept of fourth-party logistics (4PL) has been put forward for many years. But the essays on 4PL are significantly fewer than for 3PL. Some studies are trying to clarify the definition of 4PL or putting 4PL into the application. In 1996, Accenture first proposed the definition of 4PL, putting forward that “the 4PL is an integrator that assembles its own resources, capabilities and technology and those of other service providers to design and manage complex supply chains”\(^14\). Nowadays, it is widely accepted that 4PL is a certain evolution of 3PL\(^15\) as it extends the ability to integrate the entire supply chain. Thus, 4PL has become an inevitable tendency in the logistics industry\(^16\). 4PL is not responsible for transporting the freight by itself but provides optimized solutions for customers through integrating the capabilities and resources of its partners\(^17\)\(^18\), bringing huge earnings to enterprises\(^14\).

There is a relatively small body of literature that is concerned with the combination of 4PL and intermodal transport. Some papers focus on the model optimization of the intermodal transport system in which 4PL takes participates. Mes and Iacob\(^19\) model the synchronodal planning of intermodal transport and propose an algorithm to select the best combinations of modes for every order\(^19\). Van Heeswijk et al.\(^20\) consider the case that 4PL takes responsibility for matching orders to carriers and set main arcs and free arcs to solve the problem\(^20\). Sitek and Wikarek\(^15\) design the supply chain from the view of 3PL/4PL/5PL and propose a mixed integer linear programming problem to optimize the cost of production, transport, distribution and
environmental protection. However, few have studied the way of improvement of intermodal transport system from the perspective of 4PL.

Consequently, this paper aims to explore suggestions for the development of intermodal transport with the support of 4PL. In addition, the paper attempts to focus on the Yangtze River Economic Belt to make sure that the solutions are in touch with reality. The remaining part of the paper proceeds as follows. Chapter three of this paper will introduce the intermodal transport system. Based on the system, the paper identifies some problems in Yangtze River Economic Belt in Chapter four. Furthermore, Chapter five clarifies the advantages of 4PL applied in the intermodal chain. Ultimately, we propose some advice on the development of intermodal transport in Yangtze River Economic Belt from the perspective of 4PL.

3. INTERMODAL TRANSPORT SYSTEM

The intermodal transport system consists of several components and there is a need for management of the intermodal chain to coordinate different components. It is important but difficult to establish an integrated intermodal transport system.

3.1 Components

Policies. The policy of the government is thought of as a key component in the intermodal transport system. The government should clarify the strategic positioning of intermodal transport and formulates some policies to provide support such as the land policy, the logistics policy, the infrastructure policy and the environmental policy. Additionally, the government needs to make some regulations, concerned with the administration of the freight, the right and obligation of the actors, and the content and the validity of the transport bills. The Ministry of Transport in China promulgated Terminology for Intermodal Freight Transportation and Intermodal Loading Unit Marking in 2017, which filled the gap in the intermodal transport standard in China.

Actors. The intermodal transport system is composed of many different companies and related actors, including shippers, carriers, facility and physical infrastructure managers, institutional authorities and customers. Shippers generate the transport demand while carriers perform the transport. Facility and physical infrastructure managers take responsibility for the construction and maintenance of the infrastructure. Institutional authorities are the actors who formulate laws and regulations. Customers receive the freight through intermodal transport.

Terminals. An intermodal freight terminal is the place where freight transshipments between variable transport modes happen. The terminal can be the railway station, port terminal, road cargo station or airport. “Vehicle loading and unloading, cargo and vehicle sorting and consolidation, convoy make up and break down, and vehicle transfer between services” are the major operations in terminal. The efficiency and effectiveness of services in the terminal are crucial to the quality and cost of the intermodal chain.

Transshipment services. There is different equipment used for freight transshipment (shown in Figure 2). The first is loading units which mainly include semitrailers, swap bodies, and containers. Carrying tools, such as flatcars and ro-ro ships, serve as the facilities that hold the loading units. Transshipment facilities mean a lot to the mechanization of intermodal transport operations, which include gantry cranes, reach stackers and straddle carriers.
Information systems. Information interchange among variable transport modes is the essential basis of intermodal transport operations. In addition, actors in the chain of intermodal transport also need to exchange information and data[27]. The construction of the information system depends on the development of Information and Communication Technologies (ICT), which increase data flow and quality of information, improving the methods to solve operation problems in intermodal transport.[28]

Figure 2. Freight transshipment equipment
3.2 Intermodal Chain

The intermodal chain achieves door-to-door service and takes place over long distances\cite{24}, which
needs the coordination of different actors and transport modes. Commonly, shippers may be the
producers of freight or some agency companies who generate demand. Then carriers, such as
railway, road transport companies and shipping companies, take responsibility for providing
freight services. Sometimes, the freight needs to be transshipped among different modes, thus
requiring an integrated information system. Taking sea-rail-road transport for example, before
the containers arrive at the seaport, the coordinator should notice the railway to make sure the
swift transshipment. Then, the containers are loaded into the rail using the standard equipment
and another transshipment is made in the rail terminal. Ultimately, the freight is picked up by the
cargo and transported to the customers. It needs to be mentioned that the institutional authorities
play an important role in the process as the laws and regulations make sure the successful flow
of freight in the chain. Generally, involving so many market players and institutional authorities,
it’s difficult to coordinate the intermodal transport chain to some extent.

4. PROBLEMS OF INTERMODAL TRANSPORT SYSTEM IN
   YANGTZE RIVER ECONOMIC BELT

This chapter identifies problems of the intermodal transport system in Yangtze River Economic
Belt. In general, the coordination difficulties of the intermodal chain are the main barriers to the
development of intermodal transport in the region.
4.1 Policies

Yangtze River Economic Belt lacks an integrated department for intermodal transport. With many interest groups, it is still a complex task to manage intermodal transport. Thus, the government in the United States has established the National Intermodal Transport Committee to guide the development of intermodal transport at the national level[3]. However, China has not established a department or committee which works out to coordinate and resolve the hidden conflicts, resulting in many problems. For example, neighboring ports may repeatedly increase their transport recourses, leading to overcapacity[29].

From the perspective of the specific policies, although the government has promulgated many technical standards, there is no unified standard system. China has enacted more than thirty intermodal transport standards but the standard system has not been constructed yet. Thus, the standards used in the economic belt are different[5], leading to high transport costs. While the United States and Europe have established more perfect standard systems after more than a half-century of development[30].

4.2 Actors

Cross-regional coordination mechanism of Yangtze River Economic Belt has not been established. There still exist many administrative barriers to promoting intermodal transport among eleven provinces. The differences in opinions and positions make it difficult for cross-regional coordination[31]. For example, waterway regulation requires the joint decision of multiple provinces while different opinions of related provinces add to the difficulty of the regulation[32].

In addition, Yangtze River Economic Belt also lacks a unified intermodal transport plan. Different institutional authorities have little communication and information interchange, resulting in difficulty in customs clearance sometimes. For instance, institutional authorities, such as the quarantine department, customs and immigration checkpoint, repeatedly check the freight. Therefore, extra operations in transshipment between different modes result in the transport time extension[33], which needs some measures to reduce processing time[34]. Uiwang Inland Container Depot (ICD) in Korea has adopted a public-private partnership pattern. Apart from the freight companies, customs and some other government departments also lie in ICD[21], which promotes information interchange and reduces transport time.

4.3 Terminals

There exists the “middle one-kilometer dilemma” in the connection of different intermodal transport terminals. The volume of container transport increases greatly in recent years, leading to demand growth in inland transport. However, the shortage of connection resources happens between ports and inland areas, resulting in the bottleneck of intermodal transport[35]. Take sea-rail mode for example, railway connection stations in Yangtze River Economic Belt are generally constructed in the early time, which are unable to meet the increasing distribution demand of ports[32]. Some ports even have no linked railway to transport the freight and heavily rely on road transport. Among ports in more than 2,800 kilometers of the waterway, there are less than 10 ports that can carry out intermodal transport between rail and water[36]. Consequently, the volume of sea-rail transport only accounts for 1.5% of the total port freight volume in China while the proportions in the United States and Germany are more than 20%(37). It is worth mentioning that the United States has built more than 1,400 intermodal transport channels.
4.4 Transshipment Service

Mismatch of the transport equipment makes it difficult to ensure a fast transshipment process between different modes. Transshipment refers to the operation that transfers the flow from one mode to another mode[38]. However, the mismatch between different equipment, usually resulted from the inconsistent standard of equipment, increases to the difficulty of transshipment sometimes. Zhang et al.[39] point out that the pallets mismatch the containers and the loading units mismatch the flatcars in China[39]. However, the United States and Europe put great emphasis on the standardization of transport equipment. For example, the National Intermodal Transport Committee in the United States has studied the standards of intermodal transport by investigating the comparative advantages of different standards for many years[3].

The lack of “one bill of lading” is the main barrier to intermodal transport in Yangtze River Economic Belt. When the freight arrives at the terminal, it takes more time to type the information into the system repeatedly, increasing the waiting time. Additionally, the transport bill is connected closed to the transport responsibility[40]. Thus, a unified transport bill is beneficial to clarify the responsibilities among variable carriers. However, China still has no unified transport bill while 90% of the contents of different transport bills are almost the same[22]. Furthermore, some applied transport bills used in small range just suit international transport but not domestic transport[22].

4.5 Information systems

Information silos about the management of intermodal transport are common phenomena existing in Yangtze River Economic Belt. In particular, the lack of interchange of information between railway and port makes it impossible to dynamically supervise the state of in-transit freight[29]. Consequently, data accessibility which extends across various actors in the chain plays a significant role in intermodal transport[41]. However, many pieces of literature state that actors are unwilling to share valuable information due to competition[42] or security risk of community technology[43].

Some information platforms have been constructed in Yangtze River Economic Belt but the entire operation level needs to be improved[5]. The platforms only serve for sharing basic public information or transaction between different market actors. Thus, the intelligent transportation system needs to be employed to improve the operation ability in intermodal transport, such as responsiveness[44] and service quality[45]. New technologies and analytical tools are required to optimize transportation decisions[46], which include routing, scheduling, monitoring and tracking of goods[47].

5. ADVANTAGES OF 4PL APPLIED IN INTERMODAL TRANSPORTATION

A single 3PL is unable to provide integrated service for the intermodal chain, while 4PL can be the leader of the entire supply chain that integrates many 3PLs and other service providers such as consulting, financial and IT departments. 4PL integrates almost all companies involved in the intermodal chain[48]. In practice, 4PLs mainly have consulting companies, IT companies, 3PLs and logistics platforms led by government. Commonly, government establishes 4PL platform
with the support of professional private companies, which is beneficial to the development of intermodal transport.

![Diagram](image)

**Figure 4.** Advantages of 4PL in the coordination of intermodal chain

Based on integrated 4PL platform, the information interchange of different actors and modes in the chain is improved. The rise of 4PLs promotes the development of the network model which is replacing the traditional linear model in the supply chain[49]. That’s because 4PL creates the opportunity for information integration through the application of IT systems, communication flows and so on. 4PL combines the customers with other actors in the intermodal chain[50]. In addition, 4PL promotes the information interchange among rail, port, road and air, thus promoting the "one bill of lading" and reducing the time on typing the information into another
freight management system repeatedly. Information interchange and integration lay the foundation for improving the operation of intermodal transport.

4PL achieves the integrated design of the intermodal transport system. It is regarded as a point of the interface in the intermodal chain[18], which promotes the match of supply and demand. 4PL integrates a large quantity of 3PLs[51] and makes the advantages of existing operation resources of firms to meet the demand[52]. Furthermore, it not only considers the logistics decisions, such as vehicle routing and scheduling but also pays attention to solutions for the management of infrastructure, financial transactions, human resources and so on[46]. Thus, 4PL can improve almost all activities in the entire intermodal chain.

4PL also provides professional management knowledge and rich experience for intermodal transport enterprises. Most logistics firms and 3PLs are small and medium-sized enterprises and put more emphasis on scale expansion but not on improvement of service quality. Additionally, it is uneconomical for them to form an intelligent management system to improve the management level only for themselves[53]. While 4PL has an edge in formulating innovative and professional solutions as it integrates the abilities of lead consulting companies and IT providers[54]. What’s more, the abilities can be shared by varieties of companies in the intermodal chain which splits the cost. Nowadays, 4PLs are trying to improve their operational level in the hard environment (i.e. material flow, capital flow and information flow) and the soft environment (i.e. organization style and operational procedures)[16].

Finally, 4PL can play the role of coordinator between the government and logistics firms. As 4PL collects most of the data and information in the intermodal chain, it provides powerful support for government decisions. For example, 4PL can identify the bottleneck of the intermodal chain through data analysis and provide advice for the government in infrastructure construction. Especially for the Public-Private Partnership logistics platforms, the government has more access to obtain the operation data and find out the problems.

6. SUGGESTIONS FOR 4PL APPLIED IN INTERMODAL TRANSPORTATION

The design of the information system of 4PL is first considered in this chapter. Then this part provides more advice on the information standardization, coordination, utilization and innovation from the perspective of 4PL. In addition, China government needs to channel great effort into the development of intermodal transport.
6.1 Design of information system

4PL is considered the coordinator among different actors and it must obtain detailed information to improve efficiency in the supply chain. Information management is crucial to promoting the service level of the 4PL platform. The features of the 4PL platform can be identified to clarify the role of 4PL in the intermodal chain. This part also constructs the framework of the 4PL platform based on BI-lateral Resource Integration Service (BIRIS). The conception of BIRIS is first put forward by Wang and Xu[55] and has been applied to many fields. This part considers the intermodal transport from the perspective of BI-lateral market and creatively designs the framework which regards the 4PL as the coordinator of BI-lateral market.
Figure 7. Service framework of the 4PL platform
Figure 8. Technological framework of the 4PL platform based on BIRIS

Figure 9. Technological components of the 4PL platform
The features of 4PL platform include infrastructure, info-mediation, info-culture and info-structure [18] (shown in Figure 6). The 4PL platform is built on the infrastructure and it should gather various leading enterprises and professional staff in different fields, which develops an entry barrier to the 4PL industry. Info-mediation refers to various software, such as transaction software and schedule software, which are applied to optimize the material flow, capital flow and information flow in the intermodal chain. Info-culture means that 4PL takes part in the broadcasting of common standards and takes responsibility for controlling legal and regulatory activities. Consequently, 4PL is required to link the platform with some institutional authorities to deal with customs formalities, transport disputes and so on. Info-structure is referred to the function of process tracking in the 4PL platform. 4PL should collect information of the freight from origin to destination and track the time-table of different modes to make sure the fluent flow.

The framework of 4PL platform is constructed based on BIRIS55, which is a popular service pattern that is composed of customers, service providers and a third-party platform. Applying this pattern to the field of intermodal transport, 4PL is supposed to serve as the coordinator to connect the shippers and service providers. Figure 7 and Figure 8 respectively show the service framework and technological framework of the 4PL platform. The technological framework of the 4PL platform has five main components, including management layer, execution layer, application layer, equipment layer and interface layer (shown in Figure 9). Actors including shippers and carriers can log into the portal website to manage freight orders. 4PL can provide services such as personalized customization for shippers and transshipment scheduling for service providers through the data integration and analysis in the execution layer and application layer. The realization of the interaction in the front end is based on the infrastructure in equipment layer. Meanwhile, in interface layer, 4PL achieves integration of information through connecting existing business system of the actors into the 4PL platform.

6.2 Promotion of information utilization

Initially, the most important thing is to establish a unified information standard. As the interdependence increases among different actors, 4PL needs to unify Electronic Data Interchange (EDI) format to ensure information flow smoothly. In particular, the popularization of “one bill of lading” is crucial to improving efficiency in this data interchange system, while the government has emphasized that in the 14th Five-Year Plan. Therefore, 4PL has the responsibility to broadcast the unified transport bill parting with the government. The unified transport bill is required to adapt to different transport modes, including rail, water, air and road. Through analyzing multiple transport bills used in practice, the designer should study and formulate the transport bill which suits the entire intermodal chain. The transport bill should include not only basic items generally used but also optional items especially for some transport modes.

Based on the unified information standard, 4PL should promote information interchange and establish a communication mechanism for intermodal transport. The data from the internal system of different actors converge in the 4PL platform, preventing data silos. In addition, 4PL should provide a platform where multiple interest groups can exchange their opinions. The coordination of these groups is an essential motivation for pushing the development of intermodal transportation [56]. Therefore, with the support of 4PL, the cross-regional communication mechanism should be made available to coordinate cross-regional decisions, such as waterway
regulation. According to The Work Program for Development and Adjustment of Transport Structure (2021-2025), it is important to promote the open exchange of government data on intermodal transport. As the interest groups have been extended to the financial department, customs, inspection agencies and some other institutional authorities, the 4PL platform ought to open data interfaces to these institutional authorities to ensure smooth freight flow. Furthermore, great efforts are needed to converge various intermodal enterprises, promoting cooperation and information interchange. As a result, 4PL can align the operation processes in the intermodal chain to reduce possible problems.

There is a definite need for 4PL to improve its ability of information utilization, promoting the seamless connection between different transport modes. This requires the 4PL platform to encourage all transport modes to share information including time-table, the capacity of vehicles, location of the freight and so on, which make up of the checklists used for processing all shipments. To achieve effective utilization of information, 4PL should redesign the information system and use the algorithm to optimize some operational decisions such as vehicle routing, freight routing, infrastructure construction and so on. Furthermore, 4PL needs to facilitate the ability to design integrated solutions for the entire intermodal chain. It is required to optimize not only transport activities but also the operations of inventory control, order management, funds management and some other activities. Consequently, 4PL becomes the coordinator who controls the supply chain completely and provides customized solutions for different customers. Based on the integration of all modes and activities, 4PL can maximize the benefits of the whole supply chain.

4PL should evolve and innovate constantly through the application of advanced information technology. Information and communications technologies (ICT) are beneficial to information interchange and seamless connection in intermodal transport, which include “Internet of Things” (IoT), artificial intelligence (AI), big data analytics, cloud computing and machine learning. 4PL is supposed to drive the universal application of Radio Frequency Identification (RFID) and Global Positioning System (GPS) to track, monitor and control vehicles and freight. Meanwhile, big data and cloud computing could be applied to serve for route optimization, mode choice and some other decisions. For example, 4PL needs to bundle smaller batches of goods into large flows to achieve high-frequency transport.

6.3 Government supporting measures

The government ought to take some measures to support the development of intermodal transport, thus helping take the most advantage of 4PL. It is worthwhile to mention that the development plan of intermodal transport in Yangtze River Economic Belt should adapt to the industry development. The plan needs to consider the features of different regions such as the economic development level, the major industries and the condition of infrastructure. For example, the cities in the upper reaches of the Yangtze River continue to promote electronic and communication device manufacturing by linking CHINA RAILWAY Express to the inland areas. The cities in the middle reaches of the Yangtze River could focus on competitive industries such as the engineering industry and enhance the ability to transship different engineering equipment. The cities in the lower reaches of the Yangtze River play the role of promoting international trade. More specific suggestions are given as follows. Initially, the establishment of Yangtze River Economic Belt development committee is needed to enhance the coordination between different interest groups and promote the formulation of some related regulations. Different interest groups
are supposed to attend communication conferences every year to report problems in time and find out solutions together. Secondly, a key policy priority should be to plan for the long-term care of intermodal transport. China can formulate various preferential policies for intermodal enterprises, such as providing cheap lands and tax breaks. In addition, the special fund should be made to support the development of talents in this field. Thirdly, continued efforts are needed to improve the collection and distribution system of intermodal transport. In particular, railways should be constructed to connect the ports and the infrastructure needs to be improved to lift the capacity of inventory, handling and delivery. Finally, to ensure seamless connection, China is supposed to unify the standard of intermodal transport equipment, including loading units, carrying tools and vehicles.

7. CONCLUSION

This study sets out to explore the way of improvement for intermodal transport in Yangtze River Economic Belt. Consequently, this paper initially introduces the intermodal transport system, including policies, actors, terminals, transshipment services and information systems. Then some problems of the intermodal transport system have been identified. To find out whether 4PL suits the intermodal chain, the paper points out the advantages of the combination of 4PL with intermodal transport. Furthermore, the study tries to design 4PL platform based on BIRIS. The most important finding to emerge from this study is that 4PL can indeed promote information standardization, coordination, utilization and innovation in the intermodal chain. Meanwhile, the government is supposed to take some measures to support 4PL and other enterprises. Overall, the integrated intermodal transport system based on the coordination of multiple sectors, agencies and modes is crucial to the development of the regional economy in Yangtze River Economic Belt. It not only benefits the reduction of the cost of transport but also strengthens the links of city strips along the river and promotes three-industry integration. Thus, this paper contributes to exploring new ways of development of Yangtze River Economic Belt. The practical implication of this paper is that government can consider taking advantage of the 4PL platform for intermodal transport. Finally, considerably more work still needs to be done to establish a more detailed and professional 4PL platform for intermodal transport.

REFERENCES


Frost, S., [Frost & Sullivan Award Underscores Novamont’s Profile as Product Innovator], (2004).


Research and Application Analysis of Two-Level Operational Risk Prevention and Control System of "Online State Grid"

Yuling Chen¹*, Wei Zhang², Ying Jiang², Fei Lou², Qirui Chen²

*chenyuling_0904@foxmail.com
State Grid Hanzhou Power Supply Company, China¹
Zhejiang Marketing Service Center (Metrology Center), China²

Abstract: In order to solve the operational risks faced by the "Online State Grid" APP, this paper proposes to study and construct the operational risk prevention and control system at provincial level and headquarters level. Establish a closed-loop risk control mechanism including risk identification, risk warning, risk disposal, risk interception and risk re-evaluation, and create a two-level operation risk control process including the "Online State Grid" APP headquarters operation risk control center, 95598 customer service center, A city operation risk control center and local companies; Through the construction of operation system and the improvement of management mechanism, the research results of this paper have important guiding and reference value for establishing and perfecting the operation risk prevention and control system of "Online State Grid" APP, and comprehensively improving the active defense capability of "Online State Grid" APP.

Keywords: "Online State Network", Two-Level Operational Risk Prevention and Control, Defensive Ability, Risk Identification.

1 INTRODUCTION

"Online State Grid" APP is the official and unified online service population of State Grid Corporation of China, which integrates five service scenarios of "residence, electric vehicle, shops, enterprises and new energy" to meet customers' all-round and personalized power demand [1]. While the "Online State Grid" has achieved good operational results, it will inevitably attract the attention of a large number of black goods. By registering a large number of accounts, repeatedly binding users' numbers, and taking advantage of loopholes in activity rules, they encroach on the rights and interests of ordinary users and take a large number of activity points, resulting in the loss of users' rights and interests [2]. The operational risk prevention and control situation of "Online State Grid" is becoming increasingly severe, and there is much room for improvement in risk identification means. With the upgrading of malicious network attack technology and the updating of its means, the trend of risk prevention and control of "Online State Grid" is becoming increasingly severe. There is much room for improvement in risk closed-loop control and risk prevention and control system construction [3]. Multi-level coordination of operational risk prevention and control. How to efficiently carry out automatic, intelligent risk analysis and identification, risk early warning, risk disposal, and
follow-up management based on risk control rules, accumulate risk prevention and control experience, solidify and form standardized business norms and processes, quickly identify risks, reduce risk losses as much as possible, and comprehensively improve operational risk prevention and control capabilities are the key points of operational risk control.

2 "ONLINE STATE GRID" TWO-LEVEL OPERATIONAL RISK PREVENTION AND CONTROL SYSTEM RESEARCH

2.1 Risk Control Objects and Risk Control Rules

Based on the operation data of "Online State Grid", the operation big data warehouse is extracted and constructed, including user account data, user behavior data, activity participation data, rights and interests operation and other data, and the objects of operation risk control are discovered as shown in Table 1[4].

<table>
<thead>
<tr>
<th>target</th>
<th>explain</th>
<th>Risk characteristics</th>
<th>Wind control scene</th>
</tr>
</thead>
<tbody>
<tr>
<td>account number</td>
<td>Register an account on the Internet.</td>
<td>Virtual mobile phone number and account recommendation relationship</td>
<td>User behavior risk control, operational activity risk control and point management risk control.</td>
</tr>
<tr>
<td>Household number</td>
<td>Online national network binding account number</td>
<td>Account binding unites a large number of account numbers, account numbers in the same period, etc., and a large number of prizes are written off.</td>
<td>User behavior risk control, operational activity risk control and point management risk control.</td>
</tr>
<tr>
<td>activity</td>
<td>Operational activities</td>
<td>A large number of users participate in activities.</td>
<td>Operational risk control</td>
</tr>
<tr>
<td>prize</td>
<td>Operating prizes</td>
<td>Users get a lot of prizes, and the prizes are written off through multiple account numbers.</td>
<td>Operational risk control</td>
</tr>
<tr>
<td>Ip address</td>
<td>Account login IP address</td>
<td>Log in with a large number of accounts at the same IP address.</td>
<td>User behavior risk control and operational activity risk control</td>
</tr>
<tr>
<td>terminal device</td>
<td>Account login device</td>
<td>Log in to a large number of accounts of the same device.</td>
<td>User behavior risk control and operational activity risk control</td>
</tr>
<tr>
<td>integration</td>
<td>On-line State Network User Incentive Content</td>
<td>Accounts get a lot of points and accounts consume a lot of points.</td>
<td>User behavior risk control and point management risk control</td>
</tr>
</tbody>
</table>

Based on the online operation data of State Grid, users' behavior trajectories are extracted, such as registration, binding, participation in activities, etc., and information including people,
time, events and so on is extracted to construct events. According to the relationship between events (including time sequence, cause and effect, co-reference, subordination, etc.), different event rules are linked, as shown in Table 2 [5-6].

<table>
<thead>
<tr>
<th>rule</th>
<th>explain</th>
<th>Risk characteristics</th>
<th>Applicable scene</th>
</tr>
</thead>
<tbody>
<tr>
<td>recommend</td>
<td>Account Registration</td>
<td>Virtual cell phone number, cell phone number of the same segment, and continuous cell phone number.</td>
<td>User behavior risk control, operational activity risk control and point management risk control.</td>
</tr>
<tr>
<td>bind</td>
<td>Account number binding account number</td>
<td>Account binding unites a large number of account numbers, account numbers in the same period, etc., and a large number of prizes are written off.</td>
<td>User behavior risk control, operational activity risk control and point management risk control.</td>
</tr>
<tr>
<td>participate in activities</td>
<td>Participate in operation activities of account number.</td>
<td>A large number of users participate in activities.</td>
<td>Operational risk control</td>
</tr>
<tr>
<td>Get a prize</td>
<td>Account number to win prizes for operational activities.</td>
<td>Users get a lot of prizes.</td>
<td>Operational risk control</td>
</tr>
<tr>
<td>Write off prizes</td>
<td>The prize is written off by the account number.</td>
<td>Prizes are written off by multiple account numbers.</td>
<td>Operational risk control</td>
</tr>
<tr>
<td>IP login address</td>
<td>Account login IP address</td>
<td>Log in with a large number of accounts at the same IP address</td>
<td>User behavior risk control and operational activity risk control</td>
</tr>
<tr>
<td>Login device</td>
<td>Account login device</td>
<td>Log in to a large number of accounts of the same device.</td>
<td>User behavior risk control and operational activity risk control</td>
</tr>
<tr>
<td>Get integral</td>
<td>Get user incentive content</td>
<td>Some accounts get a lot of points.</td>
<td>User behavior risk control and point management risk control</td>
</tr>
<tr>
<td>Consumption integral</td>
<td>Account number consumption points</td>
<td>Consume a lot of points through some account numbers.</td>
<td>User behavior risk control and point management risk control</td>
</tr>
</tbody>
</table>

2.2 Online Monitoring of Operational Risks

Based on risk rules, conduct real-time monitoring and analysis on accounts, activities, points distribution and other behaviors, set key indicators and indicator thresholds for risk monitoring, conduct online risk monitoring, update the details of risk account list in real time, and export them as needed for risk analysis and risk disposal: User behavior risk monitoring mainly monitors abnormal data such as new and active users [7]; Operational risk monitoring mainly monitors abnormal data such as the increase in the number of users participating in the activity, the distribution of prize rights and interests, and the surge in the number of rights and interests
written off; Equity risk monitoring mainly monitors abnormal data such as the number of account points obtained and the number of account points written off\(^8\).

2.3 Closed-Loop Control of Operational Risks

According to the operational risk prevention and control scenario, carry out risk closed-loop management and control based on risk identification, risk early warning, risk disposal, risk interception and risk re-examination\(^{9-10}\). Risk identification: based on the knowledge map of operational data, identify the operational risks of relevant scenarios, generate a grey list, and submit risk alarms; Risk warning: process risk events and risk data, form risk warning information, and send it to operation risk control management personnel to realize risk warning; Risk disposal: the operation risk control management personnel analyze and identify the risk warning information, classify and manage the risk events, and manually identify the grey list to form a white list or a black list\(^{11}\); Risk interception: submit the blacklist to the State Grid Risk Control Center, and the State Grid Risk Control Center will freeze relevant accounts, suspend their participation in activities, and realize risk interception\(^{12}\); Risk review: analyze and summarize the causes, evolution process, identification and disposal of risks, find out the problems of risk management and control, improve the process of risk management and control, and improve the level of risk lean prevention and control\(^{13}\).

2.4 Two-Level Operational Risk Prevention and Control Process

Based on the method proposed in this paper, a two-level operational risk prevention and control system is constructed as shown in Figure 1\(^{14}\). The operation center is responsible for risk monitoring, risk identification and risk warning, formulating risk control strategies and measures for headquarters activities and distribution of points related to headquarters activities, ensuring the lowest risk of points in headquarters activities, and generating risk list, manual identification and risk re-evaluation\(^{15-19}\). The operation risk control center of the headquarters intercepts risks based on the risk list. The territorial company shall transfer the risk information and optimize the operation according to the risk interception situation of the headquarters; Make relevant customer service preparations and customer service work\(^{20}\).

![Figure 1 Operational risk prevention and control system based on operational risk control scenario](image)
3 CONCLUSION

Aiming at the operational risks existing in the process of "Online State Grid" operational risks, this paper studies and constructs two operational risk prevention and control systems at provincial level and headquarters level. Based on the research results of this paper, the Electric Power Operation Risk Control Center promptly dealt with abnormal users' illegal points collection in activities such as "sharing courtesy" and "recommending courtesy", and frozen account points in time, effectively reducing operational risk losses. Study the establishment of provincial and headquarters operational risk prevention and control systems, give full play to the advantages of the two levels of risks, improve the efficiency of risk monitoring and risk disposal, and comprehensively enhance the active prevention capability of the operational risks of "State Grid Online".

REFERENCES


Research on Rich-Club Phenomenon in Beijing Urban Area Based on Trajectory Data

Xiaotong Guo*

*Corresponding author: 20120742@bjtu.edu.cn

School of Traffic and Transportation, Beijing Jiaotong University, Beijing, China

Abstract: With the development of urbanization, urban transportation facilities have become more and more perfect, and the connection and interaction between different areas of the city have become easier. However, few studies have explored the forms of such connections and interactions between different areas of the city. Therefore, based on the GPS track data of taxis in Beijing, this study constructed directional and weighted urban travel networks, through the OD of residents' travel tracks. In addition, the forms of connections and interactions between urban areas in Beijing are investigated with the help of complex network. It is found that the relationship and interaction between urban areas in Beijing presents a typical rich-club phenomenon, and there are core members in the rich-club. Moreover, the urban areas mapped by these core members are related to the railway stations in Beijing. There are key members in the core members of the rich-club, and these key members have more interaction with other core members, and act as a bridge for other core members to connect with each other.

Keywords: Traffic Network, Trip OD, Rich-Club, GPS Data, Taxi, Region Interaction, Beijing.

1. INTRODUCTION

With the stability of urban structure, urban residents' travel has become regular gradually. It has become a hot research topic [1, 2] to explore urban structure by using the travel trajectories of urban residents. A hierarchical and highly complex urban travel network can be constructed through the travel trajectories of urban residents. Urban mobility networks can generally be grouped into two categories: (i) small mobility networks at the city and zone scale, which refers to mobility between smaller communities within a city or transboundary administrative region; (ii) large mobility networks at the regional level between cities, provinces or countries. Due to the increasing availability of large-scale data sources such as census data, high-speed rail data and social media, a lot of researches focus on analyzing the spatial patterns and connections of large mobile networks [3, 4]. Nevertheless, few studies focus on small mobile networks. The related researches can explore the connections and interactions between urban areas based on GPS track data of taxis or private cars and swipe card data of buses or subways. From this, the paper can provide guiding opinions for urban transportation planning and urban epidemic transmission modeling.

An important research area of complex networks is the application of network metrics. In this study, rich-club coefficients are used to quantify the nature of connections and interactions
between nodes. Rich-club coefficient is used to detect the oligopoly characteristics dominated by "rich" (high degree or high strength) nodes in the directed network. In complex networks, rich-club phenomenon refers to influential or prominent nodes that strongly interact with each other to form a cohesive subgroup while maintaining contact with “poor” nodes [5]. Rich-club phenomenon is reflected in many networks [3, 4, 6], which can be used to measure the rich-club phenomenon in different real networks. For example, Wei et al. [4] revealed the rich-club effect of human migration during Chinese Spring Festival based on the data of people returning home and returning to work during the Spring Festival. The results show that rich cities dominate the interconnection with non-rich cities, influence the entire network, and significantly reshape the network structure. In addition, the study reveals the importance of "rich" nodes in complex network dynamics, such as cascading failures, targeted attacks, and epidemic spread.

Rich-club phenomenon exists in large mobile networks. However, few have focused on the following two questions: (i) whether Rich-club exists in small mobile networks; (b) What are the attributes of rich-club members in small mobile networks? In order to answer these two questions, this study builds an urban travel network with directed authority based on Beijing taxi trajectory data and Beijing rasterization data, and investigates the rich-club phenomenon of inter-regional connections and interactions in Beijing and the nature of the members who constitute the rich-club.

2. METHODOLOGY

2.1 Network Construction

The travel of residents strengthens the connection between different areas of the city, and the travel volume between different areas of the city reveals the interaction between urban areas and quantifies the strength of this interaction. Therefore, this study constructs a complex network \( G_{od} = (V, E) \) based on residents' travel trajectory OD to research into the connections and interactions between different areas of the city. The node \( v_i \) in urban travel network \( G_{od} \) represents the grid point mapped by the starting point O or end point D of a resident travel trajectory. A directed connection is established between the starting point \( v_i \) of the travel trajectory and the ending point \( v_j \) of the travel trajectory, otherwise the direction is the same as the direction of the trajectory, denoted as \( e_{ij} \). The weight \( w_{ij} \) on the edge \( e_{ij} \), is equal to the number of trajectories passing through the directed edge \( e_{ij} \).

2.2 Rich-Club Coefficient

The rich club coefficient\(^7\) \( \varphi(k) \) is used to measure the phenomenon of rich clubs in complex networks. In the topological network, all nodes are ranked according to the richness (node degree or strength) parameter \( k \). The nodes whose richness is greater than \( k \) are considered to constitute the rich-club. For each club, \( r \) is the magnitude of richness, \( E_{sr} \) represents the number of edges connecting network nodes with richness greater than \( r \), \( W_{r} \) represents the sum of the weights of these edges, and \( W_{lr ank} \) represents the lth weight of network edges. \( \varphi(k) \) can be denoted as follows,

\[
\varphi_w(r) = \frac{w_{sr}}{\sum_{l=1}^{E_{sr}} W_{lr ank}}
\] (1)
Based on the existing urban travel network $G_{od}$ degree distribution, to create a random network $G_{null}$ as the benchmark. The urban travel network $G_{od}$ and the random network $G_{null}$ are calculated by Eq.(1) to obtain $\varphi(k)$ and $\varphi_{null}(k)$, respectively. The rich club effect $\rho(k)$ is obtained by the ratio of $\varphi(k)$ and $\varphi_{null}(k)$, as shown in Eq.(2).

$$\rho(k) = \frac{\varphi(k)}{\varphi_{null}(k)}$$

$\varphi_{w, null}(r)$ is the rich club coefficient obtained from the corresponding random network used for comparison. Then, the rich club effect is measured as follows:

$$\rho_w(r) = \frac{\varphi_w(r)}{\varphi_{w, null}(r)}$$

Among them, $\rho_w(r)$ greater than 1 indicates the existence of rich club characteristics. There are two ways to select a rich club node for a directed power network. Using the degree of nodes ($r = k$) to determine the rich club nodes is the first method, and the other method is based on the intensity of nodes ($r = s$). Both of them can be used to prove the importance of nodes, and these two richness parameters are used in this study.

3. DATA

3.1 Study Area

Beijing (115.7° E - 117.4° E, 39.4° N - 41.6° N), the capital of China, the political, cultural and technological center of China, is an international metropolis. As of 2020, the total area is 16,410.54 square kilometers under the jurisdiction of 16 districts. According to the research needs, 1km×1km grids are selected, and there are 16897 grids in the study area totally. Then, the passenger travel trajectory is mapped to the study area grid. Finally, the grid is used as the candidate node to construct the urban travel network.

3.2 Taxi Data

This paper uses the GPS trajectory data of taxis for five days in July 2015. Among them, each data contains 6 key attribute fields, as shown in Table 1, which record the longitude and latitude position and state of a taxi at a certain time. GPS devices in taxis record data every 2 to 4s. After data preprocessing, 235940 passenger travel tracks are obtained. The Euclidean distance from the starting point (O) to the end point (D) of the travel trajectory is taken as the passenger travel distance $d$. Considering that the travel distance is too short or too long, it has no theoretical and practical significance, therefore, this paper only keeps the travel trajectory of passengers whose travel distance $d$ is between 0.5 km and 200 km.
Table 1. Description of taxi’s GPS location data

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Sample</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>taxi license plate</td>
<td>BP**81</td>
<td>license plate number</td>
</tr>
<tr>
<td>time</td>
<td>2015/6/30 23:01:35</td>
<td>time when data is uploaded</td>
</tr>
<tr>
<td>longitude</td>
<td>116.416382</td>
<td>longitude when data is uploaded</td>
</tr>
<tr>
<td>latitude</td>
<td>39.828822</td>
<td>latitude when data is uploaded</td>
</tr>
<tr>
<td>speed</td>
<td>11</td>
<td>instantaneous velocity when data is uploaded</td>
</tr>
<tr>
<td>status</td>
<td>1</td>
<td>Zero means empty driving, and one means carrying passengers</td>
</tr>
</tbody>
</table>

4. RESULTS

4.1 Urban Travel Network

Based on the resident travel trajectory data and Beijing raster data, the urban travel network is constructed, as shown in Figure 1(a). During the construction of the network, the longitude and latitude attributes are given to the nodes of the network, which can visually see the ground connections between different urban regions in Beijing. From Figure 1(b) and (c), the distribution of node strength and the number of edges of different nodes in the network are represented, respectively. It can be seen that both of them conform to power-law distribution.

![Figure 1. Urban travel network](image)

4.2 Urban Rich-Club Phenomenon

The results of the rich club coefficient based on the node degree and the node strength are shown in Figure 2(a) and (b) respectively. As shown in Figure 2(a), the rich club based on the node degree is always greater than 1, indicating that there is a rich club in the network, and the regional tendency of the Beijing city area to be widely associated with its region is associated with similar areas. As shown in Figure 2(b), there is a general value of greater than 1 in the rich club, which is associated with a similar regional tendency and a similar region. Note that there are two great values of the curve, indicating that the rich club of the city's urban areas has a two-level structure, and the rich club has a core member.

From Figure 2(b), the value of node strength $s$ corresponding to the second extreme point is 1276. Nodes with node strength $s$ greater than 1276 are defined as core members of Beijing urban regional rich-club, and four core members are finally determined. Their grid numbers and corresponding urban areas are 9163(Beijing South Railway Station), 8798(Beijing West
Railway Station), 8269(Beijing West Railway Station) and 8976(Beijing West Railway Station), respectively. Note that the core members of Beijing Urban Regional Rich Club are all related to train stations, and the contact and interaction between train stations are relatively frequent. This may be related to the need for passengers to transfer between different stops.

As shown in Figure 2(b), when the value of node strength $s$ is greater than 1364, the node with grid number 8976(Beijing West Railway Station) exits the rich club, and the rich club coefficient between different urban regions of Beijing is less than 1, hence the rich club cannot be formed in the urban region of Beijing. Accordingly, Beijing West Railway Station is a key member of the rich club. It has more interaction with other core members and acts as a bridge for other core members to connect with each other.

![Figure 2](image)

**Figure 2.** Rich-club coefficients based on degree and strength

5. CONCLUSIONS

Firstly, this study constructs the urban travel network based on Beijing taxi trajectory data and Beijing raster data. Furthermore, the forms of connections and interactions between different urban areas in Beijing from the perspectives of degree and intensity are investigated. It can be found that there is a rich-club phenomenon in the connections and interactions between different urban areas in Beijing. From the perspective of degree, the regional tendency of the extensive connection with other regions in the urban area of Beijing is to have the connection with similar regions. From the perspective of strength, the areas with large inflow and outflow in Beijing tend to be related to similar areas.

In addition, the rich club composed of different urban areas in Beijing has a two-level structure, and there are also core members in the rich-club. Through qualitative analysis, it is found that the areas where these core members are located belong to the adjacent areas of different railway stations in Beijing, which means that residents who take taxis near railway stations are more inclined to go to another railway station, and their purpose is likely to be transfer. Among these core members, there is a key member -- the area where Beijing West Railway Station is located, which is the key to connect other core members. It can be seen that Beijing West Railway Station is the core of Beijing. Beijing West Railway Station not only has a good connection with other "poor" areas in Beijing, but also plays a key role in the connection and interaction of "rich" areas.
REFERENCES


Research on Export Tax Refund Declaration Robot Based on RPA

Yin Ping

e-mail: py153250@163.com

department College of Financial Accounting Zhejiang Institute of Economics and Trade Hangzhou, Zhejiang, China

Abstract: In recent years, with the export tax refund comprehensive service companies and higher vocational colleges to deepen cooperation, “First-year students do not learn the theory of export tax refunds, it is difficult to quickly qualified for the job; second-year students have heavy schoolwork, it is difficult to guarantee enough working hours; third-year students after just getting used to the work will face leaving the school ,” and other issues are increasingly prominent. This paper proposes to make a RPA-based export tax refund declaration robot to try to solve the above problems. This method achieves satisfactory results in practice, and compares it with the python-based export tax refund declaration robot, which shows the advantages of choosing RPA technology.

Keywords: RPA, Export Tax Refund Declaration, Robot.

1 INTRODUCTION

Both the impact of the “Belt and Road” policy and the opportunity of the "Post-epidemic era" have contributed to the rapid development of our export trade, at the same time also makes the export tax refund agent declaration demand rapid growth. The export tax refund comprehensive service companies began to cooperate with higher vocational colleges, and students completed the export tax refund agent declaration for companies every month. In the course of cooperation, the following problems appeared: first-year students do not learn the theory of export tax refunds, it is difficult to quickly qualified for the job; second-year students have heavy homework, it is difficult to guarantee enough working hours; third-year students after just getting used to the work will face leaving the school. The results of the collaborative process are not ideal.

Robotic Process Automation (RPA) is to simulate, enhance and extend the interaction process between users and computer systems according to pre-set business processing rules and operation behavior by using and understanding the enterprise's existing applications through user interface, automatic completion of a series of specific workflow and expected tasks, effective implementation of people, business and information systems integration of intelligent software (1). If RPA can be applied to export tax refund agent declaration, to complete repetitive, applicable and regular tasks, teachers and students can focus on more valuable work,
it can also increase the speed while ensuring the quality of work and laying the foundation for long-term school-enterprise partnership.

The research of RPA technology has been abundant in the field of finance and economics, but the research in the field of finance and economics is obviously more than that in the field of tax, and none of them involves the declaration of export tax refund. Based on various practical problems encountered in the agent declaration of export tax refund for enterprises in the cooperation between schools and enterprises, the design framework of RPA-based export tax refund declaration robot is proposed for the first time, the other parts of the article are as follows: the second part combs the workflow of export tax refund agent declaration and business “Pain Point”; The third part introduces the implementation process of the system; the fourth part shows the effect of the system in practice, and compares it with the robot based on python to explain the reason of choosing RPA; the fifth part is the conclusion.

2 DESIGN AND ANALYSIS OF RPA-BASED EXPORT TAX REFUND DECLARATION ROBOT

The export tax refund declaration is divided into two types: the export tax refund declaration of the manufacturing enterprise and the export tax refund declaration of the foreign trade enterprise, but the agent declaration is mainly for the foreign trade enterprise, therefore, RPA-based export tax refund declaration robot mainly based on foreign trade enterprise export tax refund declaration process to conduct a comprehensive analysis, identify the most complete business flow, lay the foundation for automated process design. The business process of export tax refund declaration of foreign trade enterprise after reconfiguration mainly includes: system login, data collection, tax refund declaration, system reinstallation.

2.1 System Login

The export tax refund agent declaration work of the foreign trade enterprise is completed by the robot automatically, but because of the design limitation of the export tax refund declaration software of the foreign trade enterprise, the business “Pain Point” of this process is: the agent declaration is divided into the first declaration and the others, the two types of declaration workflow is different, the first declaration needs to fill in the enterprise's basic information, then you can enter the business data filing, and the others can enter the business data filing directly. The enterprise's basic information filing page is shown in Figure1 below. Therefore, when designing the RPA-based export tax refund declaration robot, after each declaration, the software will be unloaded and reloaded directly, and the first declaration workflow will be adopted at the beginning of the declaration.
2.2 Data Collection

This part of the work first needs to be carried out by the export enterprises in accordance with the requirements of the export tax refund comprehensive service companies, provide the following information: enterprise basic information, customs declaration form, export agent certificate, purchase invoice, export invoice and so on. The export tax refund comprehensive service companies according to foreign trade enterprises export tax refund declaration needs, relevant data collection. It is necessary to collect a great deal of business data for export tax refund agent to declare, and the order and content of data collection are different because of the difference between enterprises and business, so the work is both tedious and time-consuming, has been considered the biggest "Pain point". But because data collection is the starting point of all agent declaration work, its accuracy often determines the accuracy of the whole agent declaration work, the staff’s financial knowledge, business familiarity, work sense of responsibility are higher, therefore, according to the content and sequence of tax refunds to produce data collection forms, RPA-based export tax refund declaration robot is able to efficiently complete the work of the necessary foundation, as follows Table 1.

<table>
<thead>
<tr>
<th>TABLE I. DATA COLLECTION TABLE OF RPA-BASED EXPORT TAX REFUND DECLARATION ROBOT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>The enterprise’s basic information</strong></td>
</tr>
<tr>
<td>Enterprise Customs Code</td>
</tr>
<tr>
<td>Enterprise full name</td>
</tr>
<tr>
<td><strong>Export information</strong></td>
</tr>
<tr>
<td>Serial number</td>
</tr>
<tr>
<td>Agency reference number</td>
</tr>
<tr>
<td>Export date</td>
</tr>
<tr>
<td>Export quantity</td>
</tr>
<tr>
<td><strong>Purchase information</strong></td>
</tr>
<tr>
<td>Correlation number</td>
</tr>
<tr>
<td>Date of purchase invoice</td>
</tr>
<tr>
<td>amount purchased</td>
</tr>
</tbody>
</table>
Where the declaration period in accordance with the export tax refund declaration requirements for collection, collection principles are: If the export date is this year, then the collection of export date; if the export date is not this year, then the collection of export date is last December.

2.3 Tax Refund Declaration

This step can be completed automatically by the RPA-based export tax refund declaration robot, after which manual review is required, it mainly compares the export tax refund amount formed by the export tax refund software with the input tax amount in the input invoice, determines whether the tax refund is complete, and considers the difference between the input tax rate and the refund rate, if you find any problems, you can check the data collection form and the declaration form generated by the system. The checking formula is as follows:

\[
\text{Purchase amount} \times \text{tax refund rate} = \text{tax refund amount}
\]

The declaration form is as follows: Figure 2, Figure 3.

2.4 System reinstallation

Foreign trade enterprises export tax refund after the installation of the declaration software can only be declared for one enterprise, if the second enterprise to declare the need to unload
reinstalled, which will not only extend the work cycle, it will also reduce work efficiency, but as mentioned earlier, the RPA-based export tax refund declaration robot will automatically install the software after each declaration, and the robot can work on its own for 7 x 24 hours, therefore, compared with manual declaration, RPA-based export tax refund declaration robot can better solve this problem.

3 THE DEVELOPMENT AND REALIZATION OF RPA-BASED EXPORT TAX REFUND DECLARATION ROBOT

3.1 System Login

Before the formal declaration, it mainly includes three steps: Login system, enterprise information collection and declaration period. The RPA design functions used mainly include: mouse click, excel cell reading, simulated keys, the detailed flow is shown in Figure 4 below. In this process, we need to set four variables: ckrq_1, ckrq_2, ckrq_3, ckrq_4.

![Figure 4. RPA-based export tax refund declaration robot login design flow chart.](image)

3.2 Data Collection

There are two kinds of data collection for export tax refund declaration of foreign trade enterprises: export detail data collection and purchase detail data collection. The RPA design functions used in export detail data collection mainly include: mouse click, excel cell read, simulated keys, trajectory moving, set variable, the number of worksheet rows, while loop, add and subtract, the detailed flow is shown in Figure 5 below. In this process, we need to set eight variables: hs_1, lv_1, ck_1, ck_2, ck_3, ck_4, ck_5, ck_6.
Figure 5. RPA-based export tax refund declaration robot export detail data collection design flow chart.

The RPA design functions used in purchase detail data collection mainly include: mouse click, excel cell read, simulated keys, trajectory moving, set variable, the number of worksheet rows, while loop, add and subtract, the detailed flow is shown in Figure 6 below. In this process, we need to set seven variables: hs_2.lv_2.jh_1.jh_2.jh_3.jh_4.jh_5.
3.3 Tax Refund Declaration

After the information is collected, we can do tax refund declaration. The RPA design functions used in tax refund declaration include: trajectory moving and mouse click, the detailed flow is shown in Figure 7 below.
3.4 System Reinstallation

After export tax refund declaration, we can do system reinstallation. The RPA design functions used in system reinstallation include: trajectory moving, mouse click, simulated keys and mouse scrolling, the detailed flow is shown in Figure 8 below.

4 APPLICATION EFFECT OF RPA-BASED EXPORT TAX REFUND DECLARATION ROBOT

4.1 Shorten the Cycle of Agent Declaration

Through the introduction of RPA technology, re-export tax refund agent declaration process, business restructuring before and after the comparison of the following figure 9.
Before using the RPA-based export tax refund declaration robot, the export tax refund agent declaration process is as follows: the original data is transmitted between the school and the enterprise; the students use the export tax refund software to declare for the first time, this process takes 1 day; teachers to review the results of declaration, because of the limited number of teachers, declaration data need to be manually compared with the original data, this process takes 2 days; according to the results of the review of students to apply for revision, because the software modification process is more complex, the process also needs 2 days; teachers to double-check, the process takes 1 day; students to double-revise, this process takes 1 day; the whole process takes 7 days. After using the RPA-based export tax refund declaration robot, the export tax refund agent declaration process is as follows: the school and the enterprise use excel to transfer declaration data; the student carries on the declaration data review, this process takes 1 day; the robot automatically declare, this process takes 1 day; the teacher reviews the result, this process takes 1 day; the whole process takes 3 days. In addition to the above reduction in the overall filing cycle from 7 days to 3 days, because filing data with preset rules was passed directly in excel and automated filing was done using RPA, put an end to the occurrence of errors in manual declaration and the application of rules, and greatly improve the efficiency of the whole agency declaration.

### 4.2 ADVANTAGES of RPA-BASED EXPORT TAX REFUND DECLARATION ROBOT

Python-based export tax refund declaration robot is far less effective by comparing with RPA-based export tax refund declaration robot. Firstly, Python uses coordinate orientation, so the software cannot run after the computer screen changes, and the universality of the robot is affected. Secondly, Python requires a high level of knowledge of computer theory, and students often just use it, cannot participate in software updates, so students can not be motivated to participate. But RPA uses visual low-code development, students can update the software directly according to their own ideas, and choose the best software by comparing the results. Finally, RPA development skills are much more transferable than Python, providing a foundation for students to use RPA technology at work in the future and enhancing their career capabilities.
5 CONCLUSIONS

The development of RPA-based export tax refund declaration robot can effectively solve all the problems encountered in the school-enterprise cooperation, and greatly improve the speed and accuracy of the export tax refund agent declaration, it also provides an effective reference for the use of RPA in the field of tax declaration in the future.

The RPA-based export tax refund declaration robot can simulate the simple and repeated operation of human beings. Most of the work in the export tax refund declaration has the characteristics of fixed flow, clear rules and high repeatability. After these links are handled by the robot, the requirements for students to master the professional knowledge are greatly reduced, first graders can also get up to speed quickly. Third-grade students do not need to spend a lot of time in familiar with the tax refund process, can be in a short time to complete multiple export tax refund experience, to achieve the goal of school-enterprise cooperation.

The RPA-based export tax refund declaration robot can work in $7 \times 24$-hour mode, has high peak processing ability, can work at any time with high intensity, and greatly reduces the working time requirement for students, second graders can do all the work without affecting the teaching schedule.

The RPA-based export tax refund declaration robot is suitable for handling large amount of error-prone business. The error-prone links are automatically completed by the robot through the rule setting, which greatly improves the accuracy of data processing. The integration and verification of data are also done automatically by the robot, which can shorten the work cycle and reduce the workload of teachers and enterprises.

On the one hand, the RPA-based export tax refund declaration robot needs to be changed according to the export tax refund policy and the change of the export tax refund software, on the other hand, it needs to be perfected continuously in the process of use, it can further improve the working efficiency of the robot and make it more humanized.

Acknowledgment: This paper is supported by "the Fundamental Research Funds for the Provincial Universities", Zhejiang Institute of Economics and Trade (Grant Number: 22QNJJ06 ).

REFERENCES


Research on the Influence of International Cooperation of ESI Highly Cited Papers Based on CNCI

Ying Tan, Yan Zhang, Mei Tang, Wanxin Zhao

Library, Northwestern Polytechnical University, Xi’an, China

Email: tangao@nwpu.edu.cn

Abstract: ESI highly cited papers represent the top level of global scientific research papers, and their influence also represents the scientific research strength of institutions. Taking the highly cited papers of Northwest Polytechnical University in May 2022 as an example, this paper analyzes the distribution of countries, institutions and disciplines of international cooperation papers in highly cited papers of ESI based on the CNCI index, explores the forms and characteristics of international cooperation in highly influential papers, and provides reference for establishing high-quality research partnership.

Keywords: Category Normalized Citation Impact, CNCI, ESI Highly Cited Papers, International Collaboration, Influence.

1 INTRODUCTION

With trending scientific statistics drawn from more than 12 million articles from over 12,000 global journals, the Essential Science Indicators (ESI) database provide comprehensive coverage for research performance benchmark and analysis. Based on a 10-year rolling file, Essential Science Indicators is updated every two months. The ESI highly cited papers are publications rank amongst the top 1% on citation counts compared with other papers published in the same field in the same year. They effectively represent the academic achievement and global impact of an research university. Since ESI highly cited papers often exhibit cross institutional collaborations from multiple disciplinaries, an emerging research topic is to characterize the international collaboration and partnership of those papers.

Recent works [1-2] leverages bibliometrics to characterize the state-of-art in international collaborations. They provide objective evaluation, but emphasized mostly on Global [3] or a country/region [4] or a specific discipline [5-6], hence failed to capture the diverse institutional partnerships. Meanwhile, most related researches have adopted criterial such as average citing rate [3,7] or H-Index [8] to represent the impact factor of an international collaboration. But these criterial also suffers from bias since citation rates vary by field and older papers are cited more than recent papers.

*The work was supported by the Development Strategy Research Fund of Northwestern Polytechnical University(No. 2022FZZ11),
Category Normalized Citation Impact (CNCI) is a valuable and unbiased indicator of impact irrespective of age, subject focus of document type. In this work, we use CNCI to quantitatively analyze the ESI highly cited papers published by Northwestern Polytechnical University (NWPU) in May 2022. We deep dived into the characteristics of the participating countries, institutions, and ESI disciplines, to reveal the principles of an influential international collaboration. Our assessment showcased the research status, strength and weakness of NWPU among its international peers. This work aims to provide educated insights for senior leaderships on resource planning and perspective partnership fostering, in hope that NWPU could further advance on the global competitiveness and research impact.

2 DATASET OVERVIEW

2.1 Data Source

We retrieved the top 1% highly cited papers from ESI database (most recently updated on May 2022) on June 12th, 2022, of which are published by NWPU between its 10-year rolling window from January 1st of 2012 to February 28th of 2022.

2.2 Methodology

We first inquires the raw ESI highly cited paper dataset from Web of Science (WOS) database and saved as Excel format, then exported them into Clarivate’s InCites analytics platform. Next, we parse the institution, region, subject information from InCites and transformed them into analytical attributes such as publication counts, percentage of international collaborated works, and CNCI values.

Note that we use both institutional name and its physical address to distinguish a unique record. During the data cleansing, we have identified and manually combined some records when a variant of institutional name present that matches with existing record. For example, the record {Nanyang Technological University & National Institute of Education (NIE) Singapore} will be merged with {Nanyang Technological University}.

Using the Excel built-in statistic functions, we can pre-process the WOS raw data and parse the collaborators types (i.e., intra-university vs. domestic vs. international) and their statistics. All of the authors / institutions in a paper are weighted equally during our analysis, meaning we count by their appearance, not by the particular order their name presents. For example, if a paper have 3 authors from 2 independent institutes, each authors and institutes gets 1 point. This method are compliant with the principles of unbiased analytics released by ESI and InCites.

3 INTERNATIONAL RESEARCH COLLABORATIONS AND PARTNERSHIPS CHARACTERIZATION

3.1 Anatomy of NWPU Published ESI Highly Cited Papers

Upon our most recent retrieval in May 2022, there are 622 highly cited papers from ESI database are associated with NWPU. The corresponding institutional CNCI value is 13.02, which far exceed the baseline value 1 and global average institutional CNCI value of 1.29, indicating
NWPU is very influential and well positioned among its international peers. Figure 1 portrays the annually ESI highly cited papers count, CNCI trends, and the ratio of international collaborations from year 2012 to 2022. Starting from 2018, the number of NWPU associated ESI highly cited papers have growing rapidly, and peak at 154 papers in 2021. CNCI impact factor during the same period remains relatively flat (except current year 2022) with the low of 7.19 in 2013 and high of 15.0 in 2021. Overall, the ratio of international collaborated papers are trending downwards, with a small rebound in 2022.

![Fig 1. Chronological distribution of ESI highly cited papers](image)

We’ve labeled the collaboration type of academic entities into: {international collaboration, domestic collaboration, intra-institute collaboration, and no collaboration}. As illustrated in Figure 2, 270 out of 622 or 43.41% highly cited papers are from international collaborations, higher than the domestic collaborations 25.21%, and intra-university collaborations 20.74%. However, the CNCI value of international collaborated papers are among the lowest (i.e., 12.08) in all categories, even lower than NWPU’s institutional CNCI value 13.02. It is clear that for NWPU, non-international collaborated papers are more influential.

![Fig 2. Categories of Collaborations](image)
3.2 The Nationality Characteristic of International Collaborators

The paper’s international co-authorship implies there are more than one authors and they are from more than one countries /regions [6]. As more authors from different institutions / countries collaborate on a paper, they expand and deepen the international co-authorship. The breadth and depth of an paper’s co-authorship usually reflect the degree of influence of this research on international stage [9].

3.2.1 Co-Authors’ Demography

We have identified 50 countries or regions (including Hong Kong, Macau, and Taiwan as regions of China) among 270 internationally collaborated papers. Figure 3 rank the top 11 countries/regions in descending order with their corresponding CNCI impact factor. Overall, each countries has more than 15 coauthored papers with CNCI value higher than 10. Collaborations between NWPU and USA (104) are among the highest, but the respective CNCI value (12.05) are unmatched for the top 3 countries in terms of CNCI value, where Canada is 14.90, Japan is 14.85, and Italy is 14.41. Whereas the collaborations between NWPU and Singapore ranked the 3rd on publications with relatively high CNCI values of 13.65.

![Fig 3. TOP11 Countries and Regions of the International Collaborators](image)

3.2.2 Collaboration Types

Table 1 presents the types of collaborations (number of countries in a coauthored paper) we identified from the 270 internationally collaborated papers. Most frequent types of collaboration are unilateral, meaning two countries or regions coauthoring a paper. Multilateral collaboration where there are more than 5 countries / regions in a research partnership ranks the least frequent types of collaboration, yet their CNCI value are among the highest. It shows the importance of multilateral collaboration and extending partnership to more countries will likely result higher impact research. In our record, the highest degree of multilateral collaboration is from a partnership with more than 15 regions or foreign countries, the corresponding CNCI value is 20.07.
Table 1 Degree of Cooperation

<table>
<thead>
<tr>
<th>Collaboration Types</th>
<th>Number of Papers</th>
<th>Proportion of Papers(%)</th>
<th>CNCI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unilateral</td>
<td>168</td>
<td>62.22</td>
<td>12.06421</td>
</tr>
<tr>
<td>Bilateral</td>
<td>67</td>
<td>24.81</td>
<td>12.03428</td>
</tr>
<tr>
<td>Trilateral</td>
<td>19</td>
<td>7.04</td>
<td>10.39675</td>
</tr>
<tr>
<td>Four sided and above</td>
<td>16</td>
<td>5.93</td>
<td>14.44934</td>
</tr>
</tbody>
</table>

3.3 The Institutional Characteristic of International Collaborators

3.3.1 Statistics of the Foreign Collaboration Institutes

Among the 270 internationally collaborated papers, we have identified 347 foreign institutes, including 36 institutes contributing 4 or more papers. In these 36 frequently collaborated institutes, 11 of them are in United States; 6 of them are in France; 3 of them are in Singapore, British, Germany and Australia, respectively; 2 of them in Japan and Russia, respectively; and 1 of each from Italy, Scotland, Sweden, and Wales.

From the collaboration frequency’s perspective, NWPU have partnered with Nanyang Technological University for 24 times over the past 10 years, with a CNCI value of 14.58 ranking at the 1st on the list. The second place of the list goes to National University of Singapore (NUS) with 16 coauthored papers and CNCI value is 13.16. Compared with the University of Tennessee System (ranked 3rd place on our list), the coauthored papers with UDICE-French Research Universities have better influential result and a higher CNCI value of 12.01, despite having fewer coauthored papers. Other 2 institutes tied on 4th place of our list are State University of New York System (CNCI is 12.47) and Moscow Institute of Physics & Technology (CNCI is 11.17). More statistics of aforementioned partnership institutes and subsidiaries are highlighted in Table 2.

<table>
<thead>
<tr>
<th>Institution Name</th>
<th>International Cooperation Papers</th>
<th>CNCI</th>
<th>Cooperative Secondary Unit(Papers)</th>
<th>ESI subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nanyang Technological University</td>
<td>24</td>
<td>14.58</td>
<td>Sch Aerosp &amp; Mech Engn(7); Sch Phys &amp; Math Sci, Div Phys &amp; Appl Phys(5); Sch Chem &amp; Biomed Engn(3); Sch Mat Sci &amp; Engn(3)</td>
<td>Material Science, Chemistry</td>
</tr>
<tr>
<td>National University of Singapore</td>
<td>16</td>
<td>13.16</td>
<td>Dept Elect &amp; Comp Engn(6); Dept Chem(5); Dept Mat Sci &amp; Engn(3)</td>
<td>Physics, computer science</td>
</tr>
<tr>
<td>University of Tennessee System</td>
<td>14</td>
<td>10.75</td>
<td>Dept Chem Biomol Engn, Integrated Composites Lab(12)</td>
<td>Material Science, Chemistry</td>
</tr>
</tbody>
</table>
3.3.2 The Number of Co-Authoring Institutions

Table 3 shows the collection of coauthoring institutions in each of the 270 internationally collaborated papers. The partnership of NWPU with 3~4 foreign institutions have published more ESI highly cited papers than any other size of partnership. When there are 4 partnering institutions, the publications CNCI value are relatively high (13.30). Whereas the CNCI value peaks at 15.03 when there are 8 or more collaborating institutes. In our record, there is a paper with largest number of partnering institutions (49) coauthored by 92 researchers from 15 countries, including United States, British, Germany, Russia. This paper have a relatively high CNCI of 20.07.

<table>
<thead>
<tr>
<th>Number of Cooperative Institutions</th>
<th>Number of Paper</th>
<th>Proportion of Papers (%)</th>
<th>CNCI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two</td>
<td>43</td>
<td>15.93%</td>
<td>13.08592</td>
</tr>
<tr>
<td>Three</td>
<td>67</td>
<td>24.81%</td>
<td>10.19134</td>
</tr>
<tr>
<td>Four</td>
<td>61</td>
<td>22.59%</td>
<td>13.3002</td>
</tr>
<tr>
<td>Five</td>
<td>39</td>
<td>14.44%</td>
<td>11.7874</td>
</tr>
<tr>
<td>Six</td>
<td>20</td>
<td>7.41%</td>
<td>12.81043</td>
</tr>
<tr>
<td>Seven</td>
<td>17</td>
<td>6.30%</td>
<td>8.4299</td>
</tr>
<tr>
<td>Eight or more</td>
<td>23</td>
<td>8.52%</td>
<td>15.03308</td>
</tr>
</tbody>
</table>

3.4 The Disciplinary Characteristic of International Collaborators

All 270 internationally collaborated papers are from 16 out of 22 ESI disciplines. As illustrated in Figure 4, Material Science and Engineering having the largest collection of publications with international collaborators, although their CNCI value are not the highest. Mathematics and Physics having the highest CNCI value of 16.54 and 15.42, respectively, followed by Geoscience (14.80) and Clinical Medicine (14.05). Note that these disciplines happen to be the ones that ranked top 1% among all NWPU’s discipline. Another coauthored paper between NWPU and NUS in the field of Social Science, has becomes quite influential with CNCI value of 10.56. This paper has contributed considerably in helping NWPU’s Social Science becomes top 1% discipline. To summarize, expanding international partnership and
collaborations can improve the research impact of NWPU significantly, especially for disciplines with less degree of internationalization such as Social Science.

![Fig 4. Disciplinary Statistics of International Collaborations](image)

We also cross-referenced the collaborating institutes and their regions with following observations: NWPU had a close partnership with institutes in United States, Saudi Arabia in the field of Mathematics; for Physics, NWPU had more collaborations with institutes in Singapore, United States, British, and Italy, especially had higher influential works with NTU and NUS; in the field of Clinical Medicine, NWPU had often published with University of Sydney and University of Adelaide in Australia, their high quality works all exceed CNCI value of 12. We should consider expanding our partnership with these institutes in established disciplines, while granting additional resources when recruiting talents from them.

### 4 CONCLUSIONS AND REVELATIONS

This paper presented a bibliometrics analyze based on CNCI and investigate the impact factors of NWPU’s ESI highly cited papers (internationally collaborated portion) from regions, institutions, and ESI disciplines perspective. We have extrapolate the following characteristics from this research:

1. The number of ESI highly cited papers are growing over years, with a relatively steady trend on CNCI values, but the percentage of internationally collaborated papers are decreasing. Meanwhile, internationally collaborated papers have a smaller fraction (43.41%) in overall publications than domestic collaboration, but also exhibit lower CNCI values. Both domestic and internal collaborations have higher CNCI values.

2. Among the 50 countries that NWPU collaborated with, United States ranks the most frequent collaborator, but the coauthored works’ impact factors are unmatched with collaborations with Canada, Singapore, Japan and Italy. In addition, NWPU’s collaboration breadth and scale still
have room to improve. We observed the research impact grow with the size of partnership, but currently NWPU still have a large number of unilateral collaborations.

(3) We have identified 347 collaborating institutes, out of which 36 have coauthored more than 4 papers. NWPU and NTU have the largest number of coauthored papers, whereas collaborations with NUS, New York State System, Moscow Institute of Physics & Technology yields high influential papers. Among all internationally collaborated papers, partnership among 3–4 institutes have published more papers than any other size of partnerships. Moreover, collaborations among 4 institutes has the highest CNCI value (13.30) than others. But in general, a paper’s impact factor grow with the size of collaborations.

(4) We have identified 16 ESI disciplines among all internationally collaborated papers. Material Science and Engineering having the largest publication base, followed by other highly influential works from Mathematics, Physics, Geoscience and Clinical Medicine. We saw diverged disciplinary characteristics in different countries/regions and institutional variety.

To conclude, NWPU still lack on breadth and depth of international collaborations and partnership. Comparing with the top-tier research universities around the globe, many disciplines’ degree of internationalization still have rooms to improve. In the context of China’s opening up to the outside world in education and the promotion of "double first-class" construction, more collaborations with foreign institutes can significantly improve the NWPU’s degree of internationalization, and further exhibit the influential on the global stage. It also plays critical strategic roles in fostering a beneficial relationship and sharing culture with other countries/regions.

Therefore, NWPU could further expand and deepen of our collaborations with international partner, with emphasis on the quality and outcome rather than quantity of publications. We should hosting more international seminars, exploring visiting and exchange opportunities with foreign partner institutes, alignment with latest scientific trends to broaden our visibilities and upgrade our mentality, to further expand the talent pool of NWPU’s research community and improve the impact on global stage.

By fostering and establishing a better visiting scholar exchange program, we can leverage on more resources to boost our own development. And by strategically planning and allocating premium resources, we can strengthen inter-disciplinary cooperation and promoting higher degree of partnership. By exploring potent perspective partners, we can eliminating the gaps between our top disciplines with other fields, enhancing global collaborations in various domain with balanced development strategy.

REFERENCES


Evaluation Method of Post Competency Based on Fuzzy Analytic Hierarchy Process

Di Cui

e-mail: cuidi2017@126.com

State Grid Liaoning Electric Power Supply Co., Ltd, Shenyang, Liaoning, 110001, China

Abstract: In the recruitment process, enterprises will focus on the adaptability of candidates and assign posts to them according to their characteristics. However, at present, enterprises still face the problem of difficult talent team construction. Based on this point, this paper puts forward the research on talent team construction of post competency management based on fuzzy analytic hierarchy process. Firstly, competency assessment indicators are established according to different post properties. The index weight analysis based on fuzzy analytic hierarchy process is realized, and the evaluation index is assigned by the relevant units of the enterprise, and 5 employees are selected as the evaluation objects, so as to verify the evaluation method in this paper. The experimental results show that the traditional evaluation method based only on experience has obvious shortcomings and cannot effectively reflect the post competency of employees. It also puts forward some improvement strategies, such as introducing high-tech talents and providing special talent management services for enterprises, in order to provide references for the construction of enterprise talent team.

Keywords: Post competency management, Fuzzy Analytic Hierarchy Process, Talent team construction, Ability training

1 INTRODUCTION

Since Taylor, the father of management science, put forward the management competency movement, the research on human resources competency has been gradually put on the agenda and widely used in the field of management. The concept of competency was first proposed by McClelland. He clearly points out that the difference between excellent completers and average completers in a certain job reflects the difference in competency, which can be divided into different factors and can be effectively measured, counted and clearly distinguished by the individual characteristics of excellent and average performance. The concept of competency is proposed and constantly developed in management practice, and its main contents include personality, traits, motivation, values, knowledge, skills and abilities [4]. Boyatzí proposed the Onion model of competency, including two parts and three levels: difficult to cultivate and evaluate and easy to cultivate and evaluate. The core level is personality and motivation. The outer layer is attitude and values, social role and self-image. The outermost layer is easy to cultivate and evaluate the knowledge and skills. Spencer proposed the competency iceberg model, which includes the knowledge and skills above the iceberg and the role positioning and values, self-perception, qualities, and motivation below...
the iceberg. Tim Hall (2013) proposed meta-competency. He believed that meta-competency is a skill that enables individuals to acquire abilities they do not possess. Based on competency, individuals can know their own strengths and weaknesses, so that they can give full play to their abilities in specific tasks. Mount (2016) measured more than 200 enterprise managers and, after analysis, believed that the competency model of managers should include three dimensions, namely, the interpersonal relationship dimension, management ability dimension and technology dimension. Han Yang (2016) emphasized the importance and necessity of selection and training based on competency for government organization work, constructed a competency model specifically for organizational cadres, and then analyzed the content direction and focus of selection and training of organizational cadres based on this model [1].

Traditional human resource management is a kind of human resource management based on position. Competency-based human resource management, the criteria for selecting employees not only focus on explicit characteristics such as knowledge and skills, but also pay more attention to implicit characteristics such as personality and motivation, to help employees make up for their "short board" deficiencies, so that employees can highlight stronger adaptability and competitiveness in the process of business model transformation and business diversification.

At present, micro-enterprises in cities have become a core force to drive the development of urban economy and science and technology, and such enterprises mainly rely on the ability of the talent teams in the development process, so it is very important to establish a stable and good talent team. However, when such enterprises set up talent teams, they often have problems such as shortage of talent supply and instability of talent team due to imperfect talent training mechanisms and insufficient investment in human capital, which limits the development of enterprises to a certain extent and seriously leads to the failure of enterprise management [8]. Unlike large enterprises, small enterprises only introduce market mechanism when facing such problems, and it is difficult to solve these problems by simply regulating the internal human resources management mechanism.

Based on this, this paper adopts a questionnaire survey to understand the current situation of talent team construction of small and medium-sized enterprises in this city, analyzes its existing problems, and proposes solutions according to the problems. The method of network survey is adopted to investigate the staff of certified small and medium-sized science and technology enterprises in a city. According to the hierarchical structure of personnel participating in the questionnaire, the data are sorted out, and the main problems and reasons in the talent construction of small and medium-sized science and technology enterprises are analyzed. The improvement strategies are put forward from the perspective of enterprises.

2 CURRENT SITUATION OF MODERN TALENT TEAM CONSTRUCTION

2.1 Insufficient supply of talents to meet the needs of the posts

In recent years, under the influence of the market economy and employment situation, the number of small and medium-sized enterprises around the country is increasing, and with the continuous development of enterprises, their construction scale is also expanding. Based on
this, the talent gap of such enterprises is very large, that is, the total demand for talents continues to rise. But the number of talents that meet the needs of enterprises has not increased, so there is a contradiction between supply and demand [2]. At present, the problem of "insufficient number of talents and difficult recruitment" is the consensus of the human resources industry in China. According to relevant research, the employment gap of small and medium-sized enterprises in southern coastal provinces and cities is more than 15%, and even if enterprises increase their salaries by more than 20%, it is difficult to recruit suitable talents [10].

This paper, in view of this city enterprise investigation process, from the local enterprise human resources department has also understood that it also has the recruitment difficult problem, but does not exist the labor force serious shortage phenomenon. It still has a big difficult problem, which affects the development of small and medium-sized enterprises because undertaking the personnel recruitment enterprise human resources department work pressure is still big. According to the staff of the human resources department of small and medium-sized enterprises, at present, the most scarce talents in such enterprises are mainly divided into two categories, one is high-tech talents, the other is high-level R & D talents. In terms of the shortage of quantity, the gap of high-tech talents in enterprises is larger, but in terms of the importance of enterprises, the shortage of high-level R & D talents has a greater impact on enterprises.

2.2 The poor stability of talent team construction

At present, most small and medium-sized enterprises generally have the problem of poor stability in the construction of talented team. Frequent job-hopping of employees in micro and small enterprises generally occurs in the early stage and growth stage of enterprise development. When the enterprise enters the growth stage, the human resources of the enterprise will become more stable [11]. Further analysis shows that this is mainly due to the lack of human resources training and the imperfect promotion mechanism in the early stage of development. The weak ability of talent team construction, which leads to the instability of the talent team and the outflow of talents. In the process of development, enterprises also have some concerns about brain drain and are unwilling to invest too much resources in staff training, which further leads to the decline of employee satisfaction with enterprises and further promotes brain drain [3]. In view of this, such enterprises are in sharp contrast to those with scientific promotion mechanisms. The questionnaire designed this time also investigates the reasons why employees choose to change jobs. The results of the questionnaire are shown in Figure 1.
Figure 1 Investigation of reasons for employee job-hopping

As shown in Figure 1, 68.1% of employees choose to change jobs because they want to pursue higher income, followed by about 42.1% of employees who choose to change jobs in order to choose a more suitable place of work, and about 25.1% of employees want to try new job challenges. Some employees were selected because they liked more stable large enterprises and institutions and were not satisfied with the existing corporate culture and interpersonal relationships, with the proportion of 23.8% and 7.7% respectively.

3 RESEARCH ON THE CONSTRUCTION OF TALENT TEAM BASED ON POST-COMPETENCY MANAGEMENT

3.1 Introduce new and high-tech talents in many aspects

In view of the problems existing in the process of talent construction in enterprises at present, considering the post-competence of enterprises, we can introduce high-tech research talents and high-level R & D talents from various aspects. The main way is through school-enterprise cooperation. In order to find talents who can adapt to the development of enterprises more quickly, enterprises should actively implement the school-enterprise cooperation policy with institutions of higher learning [9]. As the ultimate destination of talents, enterprises should actively participate in the process of personnel training, guide the training of talents in Colleges and universities, enhance their practical ability, and make the talents they train closer to the market demand. In school-enterprise cooperation, enterprises need to find schools that meet the job requirements. They accept some students for internship when they are about to graduate, give students enough time to adapt, strengthen their practical ability, so that they can gradually adapt to the work rhythm of enterprises, and then provide convenience for the construction of enterprise talent team [6]. In school-enterprise cooperation, enterprises select students who meet their needs to train, that is, to formulate the outline of students' practice, send special teachers to deal with all kinds of problems in the process of students' practice in enterprises, and evaluate their practical ability to determine whether they meet the standards of enterprise recruitment. Through school-enterprise cooperation, the employability of students can be improved. Talents that meet the needs of their jobs can be provided for enterprises, and market competitiveness can be provided for the development of enterprises [3].
3.2 Provide special talent management services

To solve the problem of brain drain caused by poor human resource management levels in small and medium-sized enterprises, it can be carried out from the perspective of strengthening special talent management services, that is, providing active services for talents in key areas such as labor disputes when necessary \[12\]. In view of this, it can be improved in three aspects. The first is to strengthen the information service of talent policy, that is, enterprises should establish a platform for talent policy exchange, problem consultation and feedback with other enterprises through the Internet, analyze the commonality of the problem of talent drain in enterprises, and discuss corresponding solutions \[5\]. Secondly, enterprises should do a good job in the management of professional title declaration of talents, give full play to the guiding and motivating role of professional title, and thoroughly implement that the declaration of the professional title of talents is not restricted by region, identity, working years and other restrictions, increase the publicity and business guidance of professional title declaration, and promote the comprehensive development of related work. Finally, technology-based small and medium-sized enterprises should take the initiative to establish standardized a labor security system and welfare treatment systems. If they have disputes with employees, they should solve labor disputes on the premise of legality, fairness, and justice.

4 POST COMPETENCY EVALUATION SYSTEM BASED ON FUZZY ANALYTIC HIERARCHY PROCESS

4.1 Establishment of post competency evaluation index system

To realize the accurate assessment and evaluation of workers in different positions, an assessment index system for the work content of workers is constructed. According to Spence and his wife's point of view, we roughly divide the competency into benchmark success and discrimination competency. Then, we establish a three-level index system based on the fuzzy analytic hierarchy process to analyze the ability needs of applied logistics talents, which are the general index layer, sub-index layer and sub-index layer. The system structure is as follows: the total index level is applied logistics talent ability demand factor A; the sub-index layer is divided into three parts: benchmark competency B1, discrimination competency B2 and service competency B3.

The index system adopts a hierarchical way to divide the index level, that is, to refine the content of the assessment index, which is to improve the assessment and evaluation index system. According to this method, a hierarchical index system is established, which is displayed in Table 1.

<table>
<thead>
<tr>
<th>Sub-indicators</th>
<th>Sub-indexes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benchmark competence B1</td>
<td>Ability to master basic knowledge of the industry C11</td>
</tr>
<tr>
<td></td>
<td>Ability to master professional knowledge of the post C12</td>
</tr>
<tr>
<td></td>
<td>Facility and equipment operational capability C13</td>
</tr>
</tbody>
</table>
Information technology application ability C14
Basic skills of enterprise operation C15
Cost awareness and data analysis capabilities C16

| Discriminative competence B2          | Sense of service C21              |
|                                     | Responsibility C22                |
|                                     | Interpersonal communication skills C23 |
|                                     | Self-confidence and self-awareness C24 |
|                                     | Team spirit C25                   |
|                                     | Personal morality and professional ethics C26 |

| Service competence B3               | Work service C31                  |
|                                     | External work C32                 |
|                                     | Social services C33               |

In this paper, the benchmark competency is divided into six parts: industry basic knowledge C11, post professional knowledge C12, facilities and equipment operation C13, information technology application C14, enterprise basic operation skills C15, cost awareness and data analysis C16.

The discriminating competence is divided into six aspects: service consciousness C21, responsibility C22, interpersonal communication ability C23, self-confidence and self-awareness C24, team spirit C25, personal morality and professional ethics C26.

The service competency is divided into three aspects: work service C31, external work C32 and social service C33.

According to the evaluation index system of post competency assessment, in order to ensure that the index system meets the evaluation needs in the actual work, after the completion of the design, the index system is filled according to the actual work content of the post. It intends to realize the comprehensive optimization and real-time update of the construction index system.

4.2 Index weight analysis based on fuzzy analytic hierarchy process

The fuzzy analytic hierarchy process is introduced to analyze the weight of different indicators to ensure the authenticity and reliability of the evaluation results. In the process of analysis, it is necessary to construct a fuzzy consistent matrix of the same level evaluation index, and grasp the importance of the index in the evaluation by comparing the same level index in pairs. In this study, the commonly used scale method is applied to compare the index weights. The scale value is between 0.1 and 0.9. The description of the fuzzy scale method is displayed in Table 2.
Table 2 Index fuzzy scale

<table>
<thead>
<tr>
<th>Scale value</th>
<th>Instructions</th>
<th>Scale definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4:0.3; 0.2:0.1</td>
<td>Assuming that the two elements to be compared are expressed as I and j, and the judgment result obtained is expressed as $r_{ij}$, the relative comparison result obtained for j can be expressed as $r_{ij} = 1 - r_{ji}$</td>
<td>Relative comparison method</td>
</tr>
<tr>
<td>0.5</td>
<td>After comparing the two indicators, the importance level of the two indicators is the same.</td>
<td>Same importance level</td>
</tr>
<tr>
<td>0.6</td>
<td>After comparison, one of the indicators is slightly more important than the other.</td>
<td>Slightly important</td>
</tr>
<tr>
<td>0.7</td>
<td>After comparison, one of the indicators is more important than the other.</td>
<td>Important</td>
</tr>
<tr>
<td>0.8</td>
<td>After comparison, one of the indicators is significantly more important than the other.</td>
<td>Significantly important</td>
</tr>
<tr>
<td>0.9</td>
<td>After comparison, one of the indicators is much more important than the other.</td>
<td>Extremely important</td>
</tr>
</tbody>
</table>

Assuming that in this process, the two evaluation indexes are respectively represented as I and j. The relationship between the two must satisfy the equation relationship $rij = 1 - r_{ji}$ during fuzzy evaluation and comparison. Therefore, R is represented as a consistency matrix, and the weight values of the assessment index elements in different index layers are calculated according to the fuzzy characteristics of the matrix. The calculation equation is expressed by equation (1):

$$W_\alpha = \frac{1}{n} - \frac{1}{2\alpha} + \frac{1}{n\alpha} \times \sum_{k > l}^n r_{ki}$$  

Where, $W_\alpha$ refers to the weight value of the corresponding index I in the evaluation; n is the index layer; $\alpha$ refers to the importance level (scale value); k refers to the number of indexes; and $r$ represents the fuzzy property. Where, the value of I satisfies $i \in \Omega$. The value of $\Omega$ is between 1 and n. The value of $\alpha$ in the equation is represented by $(n-1)/2$. According to the method, the analysis of the index weight based on the fuzzy analytic hierarchy process is completed.

4.3 Evaluation of examination results based on fuzzy evaluation matrix

After the completion of the above research, through the construction of fuzzy evaluation matrix, the results of job competency assessment are evaluated. In this process, it should be clear that the fuzzy comprehensive evaluation of employees is a comprehensive evaluation of individual behavior considering the influence of various external factors and combining with the fuzzy calculation equation. Assuming that U represents the multiple factors of the job object, U can be expressed as $\{U_1; U_2; U_3; \cdots; U_m\}$. Each evaluation index represents V in n decision forms of the object to be evaluated in its state, and V is represented as $\{V_1; V_2; V_3; \cdots; V_n\}$. It is known that the performance evaluation results of job competence are determined by n indicators, and the corresponding weight values of each indicator are different. In this case, the assignment of weights can be regarded as the previous fuzzy subset. The corresponding subset is denoted as A, then A is denoted as $\{a_1; a_2; a_3; \cdots; a_m\}$, where the value
of \( a_i \) must be a value greater than 0. Through the fuzzy evaluation analysis of each factor in the evaluation index system, a complete judgment matrix can be obtained. The judgment matrix is expressed as \( R \), and the corresponding expression of \( R \) is as follows formula (2):

\[
R = \begin{bmatrix}
R_1 \\
R_2 \\
\vdots \\
R_n
\end{bmatrix} = \begin{bmatrix}
r_{11} & r_{12} & \cdots & r_{1n} \\
r_{21} & r_{22} & \cdots & r_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
r_{m1} & r_{m2} & \cdots & r_{mn}
\end{bmatrix}
\]

(2)

Where, \( R_i \) refers to the single-factor evaluation result of the \( i \)th evaluation index, and it can be considered that the calculation factors of \( R_i \) and \( r_{ji} \) maintain a certain consistency in terms of comments. To ensure the consistency of the frequency distribution of different indexes in the evaluation process, the evaluation indexes are normalized, that is, the calculated value of \( \sum_{j=1}^{n} r_{ij} \) is equal to 1.0.

After ensuring that the relevant indicators in the evaluation process meet the requirements, the evaluation results for the work competency assessment are output by means of compound calculation. The calculation equation is as follows formula (3):

\[
B = A \cdot B = (b_1, b_2, b_3, \cdots, b_n)
\]

(3)

Where, \( b_j \) is the degree to which the evaluation object of work competency assessment has comments, that is, the membership degree of the fuzzy set. According to this method, the evaluation indexes are extracted from the membership data set. The single value calculation of the indexes is carried out according to the method proposed above. The final calculation results are given weight scores to realize the normalization of the evaluation results of different indexes. To sum up, it realizes the evaluation of assessment results based on the fuzzy evaluation matrix, and completes the research of job competency assessment evaluation system based on the fuzzy AHP.

5 PRACTICAL APPLICATION OF EVALUATION SYSTEM

The evaluation system is connected with a university to prove the effectiveness of the job competency evaluation system designed in this paper. A number of employees are chosen as the object of job competency evaluation. There are 5 employees in total, which are defined as \( J1 \sim J5 \). The working experience is 1 year, 2 years, 3 years, 3 years and 4 years respectively. The salary of the employee with the shortest working experience is the lowest, and the salary of the employee with the longest working experience is the highest, which can be seen from the order of salary. The evaluation of post competency in enterprises is mostly based on the
length of service and effective working hours, which has the problem of singleness. Therefore, the system in this paper is adopted to make a comparative analysis.

In this study, a number of relevant enterprises and departments were interviewed through enterprise visits and questionnaires. A total of 12 relevant leaders from the human resources department and the management department were invited to score the importance of the indicators through the pairwise comparison method. The weight table of the evaluation indicators was obtained through calculation (as displayed in Table 3).

<table>
<thead>
<tr>
<th>Sub-indicators</th>
<th>Weight W</th>
<th>Sub-indexes</th>
<th>Weight W_i</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1</td>
<td>0.3218</td>
<td>C11 0.1043</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>C12 0.2134</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>C13 0.1425</td>
<td></td>
</tr>
<tr>
<td>B2</td>
<td>0.4790</td>
<td>C14 0.1735</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>C15 0.2114</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>C16 0.1549</td>
<td></td>
</tr>
<tr>
<td>B3</td>
<td>0.1992</td>
<td>C21 0.1844</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>C22 0.2339</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>C23 0.1981</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>C24 0.1125</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>C25 0.1628</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>C26 0.1083</td>
<td></td>
</tr>
</tbody>
</table>

The performance evaluation results of five employees are calculated through the system in this paper, and the specific values are displayed in Table 4.

<table>
<thead>
<tr>
<th>Employees</th>
<th>J1</th>
<th>J2</th>
<th>J3</th>
<th>J4</th>
<th>J5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scores</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>24.8</td>
<td>13.5</td>
<td>13.4</td>
<td>12.2</td>
<td>20.3</td>
</tr>
<tr>
<td>B</td>
<td>6.8</td>
<td>21.3</td>
<td>7.6</td>
<td>2.8</td>
<td>22.9</td>
</tr>
<tr>
<td>C</td>
<td>7.8</td>
<td>3.5</td>
<td>18.1</td>
<td>2.5</td>
<td>19.5</td>
</tr>
<tr>
<td>Total</td>
<td>39.4</td>
<td>41.3</td>
<td>39.1</td>
<td>17.5</td>
<td>62.7</td>
</tr>
<tr>
<td>Result</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>grades</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>evaluation</td>
<td>C</td>
<td>B</td>
<td>D</td>
<td>E</td>
<td>A</td>
</tr>
</tbody>
</table>
In Table 4, among the 5 employees who participated in the post competency evaluation, the total score of J5 was higher than that of J2, J1, J3 and J4. The final results do not match the original performance evaluation results of five personnel based on work experience. It is clear that the early use of the work evaluation system only evaluates the comprehensive ability and work level from a single level, and can not evaluate the different levels of post personnel from multiple perspectives. Therefore, this system realizes the comprehensive evaluation of individual ability from different levels, which is to optimize the salary and treatment of grass-roots posts, and ensure the fairness, impartiality and comprehensiveness of individual evaluation.

6 CONCLUSION

In the process of enterprise investigation, based on the consideration of post competence, there are shortcomings and team building is not stable enough. This paper gives improvement countermeasures from two aspects. First, they should introduce high-tech talents in many ways. Enterprises can strengthen school-enterprise cooperation by participating in the training process of talents, so that they can meet the job requirements and provide convenience for the components of the enterprise talent team; Second, they should take the initiative to provide management services for special talents. Enterprises should keep abreast of the relevant national policies on talent management, and actively communicate with other enterprises.

When evaluating the professional titles of talents, they should strictly follow the relevant policies and personal abilities, so that they will not be affected by the region, age and working years. When there are labor disputes among talents, they should also strictly abide by the relevant systems to safeguard their rights. Through the above two improvement strategies, we can optimize the talent team management of enterprises from two aspects of introducing talents and retaining talents.

REFERENCES


Research on Public Participation Model Based on Block-Chain Mode: Taking Sponge City Construction as an Example

Jin Shang*

*JShang5@sheffield.ac.uk

School of Architecture, University of Sheffield, Sheffield, England

Abstract: Public participation in decision-making and design is a symbol of democracy [1]. Public participation provides a more creative and suitable scheme for urban design projects [2]. However, with the emergence of some unique plans such as sponge cities and smart city projects, design companies and government could not organize effective public participation with stakeholders, like local people or computer technology companies. Some design companies know little about computer science knowledge and the demands of local people in smart city projects. So this essay would introduce new public participation based on block-chain mode. This model allows task publishers to disassemble tasks and directly releases some professional tasks to specialized technology companies, thereby achieving more efficient project cooperation. Simultaneously, in this model, the urban designer assigns the design task of community reconstruction to the community resident group, and the resident group designs their community. To achieve the goal of maximizing public participation in community reconstruction, and make community reconstruction and urban design more democratic. What’s more, this essay would also introduce how to apply it in the sponge city construction project.

Keywords: Public Participation, Block-Chain Mode, Sponge City Construction.

1 INTRODUCTION

The definition of public participation is that citizens could directly participate in the decision-making. Public participation is a complex work and it could provide many benefits for urban design or policy-making [3]. The most important benefit is that it could provide more acceptance design schemes for local people. For example, if they could participate in the community renovation project, they could redesign their communities based on their demands. And then they could submit it to the design companies and government.

However, some projects are mainly led by the government and design companies in some countries, the public just participates in the seminar to share their ideas at the end of the project[2]. So this project would study a new model of public participation based on Block-chain mode for urban design or planning to help stakeholders have better cooperation with urban designers in urban design and planning projects.
2 CASE STUDY

The project occurred in the northern part of New England in 1988. Approximately 1 million acres of forest land were sold in the project. The project has aroused local people's attention to forestry development and use and future use rights. The local people are also worried about the unstable ecological environment and forest economy that the project may cause. Therefore, the government organized relevant departments to investigate it and collected opinions from various states, and set up a committee to take charge of this matter. In the early stage of its work, the council devoted itself to listening to and studying the different concerns, interests, and values of the residents of the area, and emphasized the development of a public participation plan. To better conduct public participation, the board of directors hired two public participation consultants to provide them with procedural suggestions. During the participation process, the council encourages the public to make suggestions and transparently handles their requests, showing the public how the council responds to various proposals. As the project progresses, the committee holds meetings with all stakeholders and collects their opinions.

This project shows a better public participation case. The committee fully collected public opinions and showed them how the committee handled these public opinions. However, it needs the committee to connect different agencies to meet the demands of the public. This process will consume a lot of workforce and material resources. What’s more, when stakeholders ask some professional questions, which committee could not respond. They have to invite experts in this field to answer this question. This essay will study a new participation model, which could connect participation with related agencies and technology companies directly.

3 METHODOLOGY

A block-chain can be thought of as a public ledger in which all committed transactions are stored in a block-chain. This chain grows as new blocks are attached to it. Block-chain technology has some key features such as decentralization, persistence, anonymity, and auditability. This essay would simulate a participation model based on block-chain mode.

There are mainly two important parts of this model. The first one is participation levels division. One project may have several participation levels. Every level has specific participation groups. The second part is the decentralization of tasks and powers. The model decomposes the project tasks. The higher-level groups will directly issue tasks and rights to the next-level group. The next level group will decompose and delegate the tasks. Finally, when the last participating level group completes the task, all tasks will be uploaded to the previous level. The highest-level group, the initial publisher of the project, will summarize and organize all completed tasks and form a complete design plan. According to figure 1, when the government wants to publish a community project, they would distribute the design tasks to the next level of participation, different design companies, to complete community renewal projects in different communities. After receiving a design task, the design company will decompose the design task and send the specific design task to the community committee. After receiving the task from design companies, the community committee divides the public space and private space of the community and distributes the task of updating these spaces to each household. In this project, the households are the last level of the project's participation level, and they renovate and design...
the public and private spaces for which they are responsible. During this period, there will be an expert consulting team to provide design and technical guidance. After these households have completed the plan design, they will upload the plan to the community committee. The neighborhood committees upload it to the design company after preliminary integration. The design company finalizes and integrates all the submitted design plans, fine-tune inappropriate plans, and submits the final renovation plan to the government.

This project model changes the urban renewal model centered on the design company and takes stakeholders as the core executors of design tasks. Therefore, stakeholders can directly participate in the design process. However, this model requires high design professionalism and a sense of responsibility for other participating levels except for the design company. Although the design company can organize an expert group to guide its work, it still needs a high level of cooperation from other participation levels to complete the task.

Figure 1: Prediction model based on block-chain mode.

4 APPLICATION

Faced with the impact of climate change on cities, some special plans such as the construction of sponge cities require the participation of the public as well as professionals from different fields to complete [6]. In the traditional project model of some countries, the design company, as the core of the project, is responsible for cooperating with other technical departments to complete the project. However, because some design companies lack knowledge in related fields such as water conservancy engineering, the design plans made by the company are unreasonable. It causes the water conservancy engineering company to be unable to perform construction. This section will introduce how to use this project model to help design companies to cooperate with other technical agencies efficiently.
Figure 2: Participation model applies in the sponge city construction.

According to figure 2, when the government publishes a sponge city plan task into a design company. The design company decomposes the construction and planning tasks of the sponge city. The design company would make a detailed description of the design task and delegate the design tasks to the water conservancy engineering company. The water conservancy engineering company is directly responsible for the design. After receiving the design task, the water conservancy engineering company decomposes the task and delegates the particular task to some specialized technical departments. For example, the geological survey task is delegated to the geological company. The site selection task is delegated to the community. The community carries out the construction intention to the residents. After the investigation, the results are reported to the water conservancy engineering company. The water conservancy engineering company sorted out the results and selected the urban areas that need to be focused on sponge city construction. When the water conservancy engineering company completes the design, it uploads the plan to the design company. The design company summarizes and finetunes the plan and then submits the final plan to the government.

It can be seen that the project model is different from the traditional planning and design project model in that it directly delegates the design tasks and rights to the relevant technical
departments, and these stakeholders are directly responsible for the relevant design. The design company only serves as a resource organizer and consolidator. This project model can improve the professionalism of the project and is more conducive to project implementation.

5 PROSPECTS AND CHALLENGES

According to figure 3, the advantage of this model is that it could directly assign some tasks to the groups or individuals most suitable for this task. For example, in a smart urban design project, the government could directly allocate the hardware and software design task of the smart city to the computer company and assign the task of urban design to the city design company. This model doesn’t need the city design company to connect with the computer company to allocate related services separately. So it saves time and eliminates some misunderstandings in task delivery. Additionally, this model could make the design more democratic. In this model, the urban designer is no longer the designer, but the assigner and coordinator of tasks. Residents and other stakeholders became designers. They can submit their ideas and plans to the design company or the government. The design company and the government integrate and adjust the plan as a whole and output the final plan. Therefore, it effectively improves the public participation part of urban design and eases the conflict between residents and the government.

![Figure 3: Task allocation model](image)

The challenge of this model is that it needs participants to have a high level of professional knowledge. For example, when community group participants in the design, they should have a basic understanding of design. To solve this problem, it requires the design company to set up some expert groups to go to the community to guide the residents to design and realize the residents’ design ideas. Therefore, this model requires a lot of workforces and material resources to achieve the final design.
6 CONCLUSION

In summary, this participation model breaks the traditional urban planning and design model. It invites stakeholders as the main designers and planners for the projects. The design company only serves as a resource summary and technical consultant in this project model. It improves the implementability and acceptability of the project. The more levels of participation are divided, the acceptability of the project plan is higher, but the difficulty of project implementation is also higher.

However, this model still requires these stakeholders to have a high sense of design responsibility and understanding of design. These issues will be improved in future related project research.

REFERENCES

Analysis of the Problems and Policy Suggestions of China's Macroeconomics Based on the Theory of Short-Term Economic Fluctuation and Numerical Simulation with Matlab

Kexin He¹, Yunyan Huang², Yujin Zhang³
452608287@qq.com¹, huangyunyan1997@163.com², 410922662@qq.com³
Hunan University of Technology, Zhuzhou 412000, Hunan, China¹
City University of Macau, Macau, 999078, China²
School of Central University of Finance and Economics, Beijing, 102206, China³

Abstract: With the arrival of the "14th Five-Year Plan", China's economy has ushered in new opportunities but it is facing great challenges at the same time. According to the short-term economic fluctuation theory, through the visual chart study of macroeconomic indicators, this paper analyzes the changing trends of consumption, investment, import and export, which are the "troika" driving the economy. Furthermore, combining the relevant macro data and theoretical model, the calibration of parameters and numerical simulation are carried out with Matlab software, and the model results support the following conclusions well that China's macroeconomics faces four problems: first, the slow recovery of consumer demand; second, the weak growth rate of the investment; third, the rising proportion of structural unemployment; fourth, the unstable situation of international trade. In the end, this paper explores the causes of these four problems and gives a series of policy recommendations.

Keywords: Macroeconomics, Chinese Economy, Short-Term Economic Fluctuation Theory, New Normal, Double Cycle.

1 INTRODUCTION

(1) Research Background

In 2021, China's epidemic prevention and control situation were good, and its economic recovery was relatively stable. Besides, China was the only country among the world's major economies which achieved the positive economic growth, and its overall situation was on the track of normalization. However, the global situation of the COVID-19 epidemic is not optimistic, and the international situation is not stable. Nowadays the world economy is complex and grim, and China's macroeconomics is also facing many new problems. At present, epidemics return frequently, and the domestic economic recovery is not stable yet; the consumption is frequently blocked, and residents' willingness to consume is declining; the investment growth is slow, and the investment structure needs to be reformed urgently; the number of structural unemployment is rising, and labor skills need to be updated urgently along with the improvement of the social security system; there are still risks in international trade and
the world economic situation is unstable. Against this background, it is of great practical significance to study and analyze the problems that China's macroeconomics is facing.

(2) Literature Review

If we want to study the overall national economic activities and aggregate analysis of a country, we can not avoid the study of this countries’ macroeconomics. Macroeconomics provides specific indicators for the analysis of the total national economy of an economy. According to it, people can judge the development of an economy by measuring these indicators, and they can use the same method to formulate development goals and plans of the economy. Much literature shows that the previous research on China's macroeconomics focuses on exploring how to implement macro-control into the development of the national economy, and studying both the differences of earnings management in different stages of the economic cycle and the means of macroeconomic governance under major public emergencies. Guo Hongmao (1994) [1] raised the question of how to implement macro-control in the development of the urban economy. The author holds that macroeconomic regulation and control must be a multi-level system dominated by central regulation and control, and he expounds his own views from three aspects of the national economic operation, regional economic level, and the relationship and difference between national and urban economic operation. In the end, he puts forward the methods to improve the level of urban macroeconomic management. Ji Yuechen (2008) [2] analyzed the evolution and existing problems of China's macroeconomic regulation and control, and put forward corresponding countermeasures and suggestions for the deviation of local governments from the regulation and control issued by the central government, and what she suggested is to make great efforts to straighten out the relationship between the central and local governments. Chen Wuchao (2013) [3] confirmed that people should consider whether the industry is cyclical or not and differences that exist in different stages of the economic cycle in the study of earnings management by setting up a model to test. Yang Zihui, Chen Yutian, and Zhang Pingmiao (2020) [4] started from the background of the COVID-19 in early 2020. Subsequently, they used the risk spillover network method to demonstrate the dynamic evolution of the risk transmission relationship among various sectors of China's financial market, and then they put forward relevant macro-governance mechanisms and risk prevention countermeasures.

In conclusion, we can find that there is a lack of analysis of problems faced by the macroeconomics under the background of the new normal in existing researches. In fact, literature that analyzes problems in macroeconomic operation does not have the current new normal background, while literature with the new normal background often only analyzes a specific industry instead of the overall macroeconomics.

(3) Methods of Study

This paper uses Keynes's short-term economic fluctuation theory, collects the basic macroeconomic indicators data of different years and different periods in China, combines with current hot policies issued by the government, visualizes the problems faced by China's macroeconomics in the form of charts, and puts forward corresponding policy recommendations in the end. This paper also combines the relevant macro data, uses the relevant macroeconomic model, carries on the calibration of parameter and numerical simulation, verifies the conclusion of the article, and predicts the future trend of the corresponding economic index.
2 CURRENT PROBLEMS FACED BY CHINA'S MACROECONOMICS

(1) The Recovery Rate of Consumer Demand is Slow and the Willingness of People to Consume is Insufficient

According to the year-on-year and month-on-month growth rate of total retail sales of consumer goods in China since August 2020 (figure 1), since the year-on-year growth rate of the total retail sales of social consumer goods reached 34.20% in April 2021, the year-on-year growth in the later period has been in a downward trend. The first reason for the decline in the growth rate is the frequent recurrence of the epidemic situation in many places. For example, in August 2021, the total retail sales of social consumption increased by 2.5% year-on-year, which is 6 percentage points down from last month due to the restriction of residents' summer travel. Second, it is affected by the lack of the national consumption willingness and the decrease of the consumption tendency. From 2011 to 2020, the growth of residents' consumption level index slowed down, and it can be seen that the growth rate of consumption level index in 2020 was -2.22%, showing a negative growth trend (figure 2).

On the basis of Keynesian theory, the higher the disposable income is, the lower the marginal propensity to consume will be. However, the traditional theory can not fully explain the current situation of China. In 2018, China's macro marginal propensity to consume was 40.5% [6], while in the same period, the macro marginal propensity to consume in developed countries such as the United States and Germany was 68.4% and 52.4% respectively. In 2019, China's macro propensity to consume was 40.4%, while this index of South Korea and Japan that are at a similar stage of development with China, is 76.2% and 76.0% respectively. Therefore, in addition to the impact of the increase in disposable income of residents, the propensity to consume of Chinese citizens is also largely affected by housing and consumption habits. As housing prices continue to rise while residents' housing demand is still strong, this contradiction leads to a condition that people reduce their daily consumption in order to increase their savings, which inhibits their willingness to consume and hinders the growth of overall consumption.

Figure 1  The year-on-year and month-on-month growth rate of total retail sales of consumer goods in China since August 2020
(2) The Domestic Economy is Transforming, the Investment Growth Rate is Slow

Under the background of "double cycle" development pattern [10], domestic economic transformation and international economic recovery, the situation of weak investment is gradually becoming an objective trend. In 2020, the growth rate of fixed assets investment in the whole society was 2.9 and it dropped dramatically this year compared with 23.6% in 2011 (figure 3). Changes in the manufacturing and real estate industries were obvious, and it can be seen in figure 4 that the cumulative growth percentage of fixed assets investment in the manufacturing industry declined month by month, which dropped from 37.3% in February 2021 to 15.7% in August. The trend shows that the manufacturing industry lacks confidence in the future market and investment growth is facing downward pressure [8].

According to figure 5, it shows that the growth rate of investment in the real estate development in China began to decline month by month after rising to 38.3% in January-February 2021, and it has fallen back to 10.9% from January to August 2021. Reasons for this decline are closely related to real estate policies issued by various regions. Nowadays, the whole country generally adheres to the basic requirement of "no speculation in housing", and some cities have introduced strict policies to curb the hot property market which has brought great pressure to real estate developers. The solvency of developers has weakened, and liquidity risk has strengthened, as a result, the housing boom index [5] has also dropped from 101.44 in February 2021 to 100.85 in August. (figure 6)
Figure 3  Fixed assets investment in the whole society (hundred million yuan)

Figure 4  Cumulative investment in fixed assets in the manufacturing industry increased

Figure 5  the growth rate of investment in the real estate development in China (%)
(3) Structural Unemployment in the Labor Market is Growing Rapidly

During the epidemic, the risk of labor supply boosted the process of replacing "human" with "machine", and more and more enterprises began to adopt industrial robots to improve production efficiency. Besides, the output growth rate of industrial robots also increased from 12.66% in 2018 to 26.86% in 2020. (figure 7)

On the one hand, the replacement of manual labor by robots has greatly improved production efficiency. For instance, the unmanned stator lamination workshop adopted by Dongfang Electric Machinery Co., Ltd. of Dongfang Electric Group has reduced labor intensity by more than 90%, and increased per capita output by 620%, and energy utilization by 56.6%. On the other hand, the popularity of robots has exacerbated structural unemployment, for that intelligent production has challenged the traditional mode of labor. In the labor market, young unemployed persons between the ages of 16 to 24 account for a large proportion of the population, (figure 8) so how to achieve full employment in the new labor market is becoming a hot topic in today's era.

![real estate climate index](image)
(4) The Export Situation of the International Trade is Unstable

With the vigorous development of international trade in the Internet era, cross-border e-commerce transactions have already become a part of people's daily life. However, in recent years, the international situation has been unstable and some kinds of big events such as the Sino-US trade war, Britain’s exit from Europe and the global spread of the new coronavirus epidemic have all indeed brought a great impact on China’s import and export trade \(^4\).

Since the reform and opening up, import and export trade has made a great contribution to China's gross national product. In 2020, the contribution rate of imports and exports of goods and services to GDP growth reached a new high of 28% in recent years. However, this mode is
facing a very serious situation when the current international trade situation is changing because China's participation in the international cycle is in the form of trade and direct investment. Once the international circulation is blocked, China may not be able to retreat and eventually return to the mode of domestic circulation \(^9\).

(5) Calibration and Numerical Simulation

In order to make the conclusions reached above more reliable, the paper establishes the corresponding macroeconomic model, and uses Matlab software for calibration of parameter and numerical simulation, finds that the model of the article can better fit the trend of several major economic indicators such as consumption, investment, export, and makes an extra-sample forecast of the future trend of several major economic indicators, and finds that in the short term, the trend of change of these indicators will not change fundamentally. For the readability of the paper, there is no specific process reported here.

3 CONCLUSIONS

In conclusion, China's macro economy is facing four problems: First, China's consumption continues to be weak due to the impact of the epidemic and changes in the consumption end of the supply chain. Second, the overall decline in confidence in the investment industry, the fluctuation of infrastructure construction and the rise of raw material prices make the overall decline in China's investment. Third, due to the combined influence of endogenous and exogenous variables, China's structural unemployment risk increases. Fourth, the sensitive international economic situation has hindered the development of international trade.

4 RELEVANT POLICY RECOMMENDATIONS

(1) Reduce Tax Reduction, Fee Reduction and Upgrade Consumption Structure and Narrow Income Gap

First, according to the tax multiplier \( k_t = \frac{\beta}{1+\beta} \) (\( \beta \) is the marginal propensity to consume), we can get the viewpoint that national income will increase with the decrease of tax revenue. At present, the starting point of income tax in China is relatively low, and the tax rate of labor remuneration is relatively high. Through tax reduction and fee reduction, the disposable income of low-income people can be increased and the marginal propensity to consume of low-income people can be increased. Consequently, the consumption demand is expected to be expanded. \[7\]

Second, improve the quality of consumption. Upgrading consumption structure, reducing purchase expenditure and medical expenditure will increase the proportion of consumption with marginal propensity to save. Based on the perfect social security of residents, it is also an important measure to promote total consumption and increase consumption growth rate.

Third, narrow the income gap between urban and rural areas, improve the social redistribution system, and promote the consumption expansion of middle-income groups. These policies can promote the stable growth of consumption.
(2) Optimize the Investment Structure and Reduce the Dependence of China's Economy on the Real Estate

First, China should make up for the shortcomings of investment in infrastructure, ecological environment protection and public safety, and expand the proportion of investment in emerging industries. Enterprises can enhance their innovation and competitiveness through investment such as intellectual property monopoly.

Second, standardize the management of the real estate industry and reduce its debt risk. The real estate investment is an important part of China's fixed asset investment that cannot be ignored, and it is the mainstay of promoting China's GDP growth. Therefore, it is significant to establish an impeccable regulation and management system for the real estate industry and to set up a special fund for stabilizing the bond market in order to avoid sharp fluctuations in the market. At the same time, reducing China's dependence on the real estate industry and reaching the transformation of economic structure and economic growth momentum should be the only way for China's economy.

(3) Create New Formats and More Jobs

First, improve the social security system and enhance national confidence in employment and life.

Second, China should promote the "expansion and upgrading" of employment and introduce preferential policies to encourage relevant employment. Establishing a complete labor training system, paying attention to adapting to the needs of the labor market, and promoting the deepening reform of higher education are practical ways. In addition, China should also popularize the re-employment training after laid-off workers, improve the quality level of workers, and let workers establish the concept of lifelong learning.

(4) Strengthen the Domestic Cycle and Further Deepen the Development of the "Double Cycle" Model

First, regard innovation and high quality as the standard, and promote the smooth flow of production, distribution, circulation, and consumption throughout the country. Besides, establish a strong domestic market, study and master relevant core technologies. To sum up, China's macro economy is facing four problems: first, China's consumption continues to be weak due to the impact of the epidemic and changes in the consumption end of the supply chain; Second, the overall decline in confidence in the investment industry, the fluctuation of infrastructure construction and the rise of raw material prices make the overall decline in China's investment; Third, due to the combined influence of endogenous and exogenous variables, China's structural unemployment risk increases; Fourth, the sensitive international economic situation has hindered the development of international trade, and promote the realization of import-free important resources and technologies.

Second, open wider to the outside world, along with strengthening the ability to deal with international risks and establishing a sound risk management system.

Third, jointly build "The Belt and Road Initiative" with high quality, improve the quality and efficiency of foreign cooperation, flexibly handle international economic relations, and promote the coordinated development of domestic and external demand.
REFERENCES

Research on the Dissemination and Evolution of Online Public Opinion on Unconventional Emergencies Based on Social Networks

Mingyu Xu

e-mail: xumingyu2021@163.com

School of Management, Shanghai University of Engineering Science, Shanghai, China

Abstract: With the advent of the Internet era, online public opinion has become a barometer reflecting social sentiment and public opinion. In the evolution of unconventional emergencies, online public opinion often catalyzes the escalation and spread of the emergencies, and even influences the development trend of public opinion. Taking the "D2809 train disaster" incident as an example, this paper constructs a social network model suitable for analyzing online public opinion on unconventional emergencies in complex systems, and analyses the role of online users in the dissemination of online public opinion, which is more conducive to grasping the evolutionary trends and movements of public opinion. In addition, this paper analyzes the trend of user sentiment evolution in public opinion networks by constructing user sentiment intensity indexes and provides a comprehensive account of sentiment changes in the process and evolution of online public opinion dissemination through visual analysis of user sentiment, to broaden a new theoretical perspective for public opinion research.

Keywords: Social Networks, Unconventional Emergencies, Public Opinion Dissemination, D2809.

1 INTRODUCTION

With the application of Internet technology in various socio-economic fields, the scale of China's Internet users continues to grow. According to the 48th statistical report of CNNIC, as of June 2021, the number of Internet users in China reached 1.011 billion, the number of mobile phone users reached 1.007 billion, and the Internet penetration rate reached 71.6%. The traditional media's one-way, delayed, and elitist communication form has been overturned by the new communication form of online media, which is interactive, instant, and grassroots and social public opinion has been projected into cyberspace and generated increasingly complex online public opinion[4]. The consequences are that they can affect public order online and offline in a short period, and even pose a threat to national security and social order. Therefore, how to deal with online emergent public events becomes a constant challenge for governments at all levels at present.

At present, scholars have made richer research findings on the characteristics of users and their interrelationships in the dissemination of online public opinion. Gong (2022)[2] and others
analyzed social and livelihood public opinion from the perspective of the evolutionary dynamics of public opinion and found that social and livelihood events have a more obvious process of online fermentation under the interaction of multiple participating subjects and that Internet users, media and opinion leaders, who exist in the exogenous dynamics of online public opinion evolution, are the driving force of online communication, with public Internet users having strong emotional demands and media and opinion leaders participating in large numbers and capable of triggering empathy psychology, which together constitutes the intrinsic mechanism leading to such events[1]. Qin (2017) et al.[5] argue that in the process of information dissemination, the general user's behavior and degree of acceptance of information is easily influenced by the attitude of opinion leaders towards an event on a hot topic, which has an important guiding role in the development of online public opinion. Naskar (2020)[3] studied opinion leaders based on expert intervention and government policies and found that expert interaction has a strong controlling effect on the spread of public opinion in social networks, and the more frequent the interaction the slower the spread of public opinion, and conversely the longer the interaction takes, the weaker the control of the spread of public opinion.

The dissemination of online public opinion on unconventional emergencies is urgent to study the characteristics of Internet users in the dissemination of online public opinion so that managers can effectively direct online public opinion on unconventional emergencies to a reasonable position promptly. Compared with other studies, this paper builds on the rich research on public opinion networks to construct a social network model for analyzing the spread of public opinion about unconventional emergencies in complex systems, to deeply analyze the basic evolution of the spread of public opinion about unconventional emergencies and calculate the metrics of user nodes to identify key users. In addition, this paper proposes a new user sentiment intensity classification metric to analyze the temporal evolution of user sentiment tendencies at different times.

2 DATA SOURCES AND PROCESSING

2.1 Data Acquisition

In this paper, we use Sina Weibo as the data source and use python's requests module package to obtain data samples related to the topic of train D2809 utilizing a web crawler. The time range selected for retrieval is: from June 4, 2022, to June 14, 2022. The specific operation steps are as follows: The headline news is selected as the starting node, and the headline news is used as an example to first crawl the D2809 train-themed microblogs related to it. The initial blog post data is a record containing fields such as the id name of the microblog, the name of the blogger, the content of the blog post, the number of likes, the number of retweets, the number of comments, the time of posting, and the time of crawling. To obtain the comments and retweets, the record is crawled again using the tweet id in the record as a springboard to obtain records containing valid fields such as the tweet id name, blogger name, blog post content, number of likes, number of retweets, number of comments, time of posting, time of crawl, etc. Cleaning of retweet comment user ids for the next sampling. After eliminating users with duplicate comments or retweets, after cleaning, there are 2223 user ids under the headline news node. 5 users are randomly selected as the second-level users under these 2223 users, and so on, and a total of 24 user nodes are obtained after finishing.
2.2 Data Pre-Processing

In the data pre-processing stage, the data is checked, exceptions are handled and the format is optimized in turn. Among them, the data checking operation is aimed at the phenomenon that the same microblog may refer to different topics, and de-duplication of the crawled data information can effectively avoid the phenomenon of data duplication. The exception processing includes processing the abnormal values of the fields of the number of retweets, comments, and likes, for example, information with 0 retweets will show "retweet" in the crawl result; information with 0 comments will show "comment" in the crawl result; information with 0 likes will show "Likes" in the crawl results, and special treatment will be given to the above anomalies. The format optimization is to adjust the value of some fields, for example, after the number of Weibo retweets, Weibo comments, and Weibo likes reaches 10,000, the Chinese character " million " will be displayed; when the posting time is close to the crawling time, it will be displayed as "Yesterday", "2 hours ago", etc. When the posting time is close to the crawl time, it will be displayed in the format of "yesterday", "2 hours ago", etc. All such data will be adjusted by the crawl time.

3 A SOCIAL NETWORK MODEL FOR THE DISSEMINATION OF ONLINE PUBLIC OPINION ON UNCONVENTIONAL EMERGENCIES

3.1 Construction of Matrix

According to the relationship between the 24 nodes collected, if there is a forwarding and commenting relationship between A and B, then the value of the relationship between them is 1; if there is no forwarding and commenting relationship between A and B, then the value of the relationship between them is 0; therefore, in the following table 1, the rows and columns represent the specific large relationship between the 24 nodes, which means that a 24x24 matrix is constructed. This matrix is then entered into UCINET to create a matrix based on the interactive forwarding or commenting relationships, as shown below.

<table>
<thead>
<tr>
<th></th>
<th>CCTV</th>
<th>dianke</th>
<th>toutiao</th>
<th>People</th>
<th>JiMU</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCTV</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>dianke</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>People</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Beijing</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>zhiqun</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ChinaNews</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>gqt</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
3.2 Visualizing The Relational Adjacency Matrix

Relationship network graphs, also known as community graphs, are now widely used in the field of social networks. Nodes and lines are the basic elements of a community graph. Nodes in a community graph represent users of the D2809 train microblogging discussion, and connections between nodes indicate the existence of retweeting or commenting relationships between users. The above Excel table was imported into UCINET, and its draw function was used to derive a social network graph of the D2809 train event community, as shown in Figure 1, where each node represents a blogger posting on the D2809 train microblog, and the connecting lines with arrows indicate the following relationship between the bloggers.

![Figure 1. Online community map of the "D2809 train incident"

According to Figure 1, we can see that the nodes at the core of the "D2809 train incident" network are people, ChinaNews, KNEWS, and gqt. This indicates that these nodes are very closely connected to other nodes, which means that they are more active in communicating with each other than other nodes. At the fringes of the network, those nodes in the social network that are not very connected to other nodes also exchange relatively little information, indicating that their influence is also small. The relationship matrix and social network community diagram above depict the connections between members of the same community. It is clear from this that there are 24 members in the community, and it is clear how they 'follow' and are 'followed' by each other.

3.3 Centrality Analysis

Centrality analysis is one of the first themes to be concentrated on in social network analysis. Centrality is used to reflect the position of social individuals in a network and what power an individual or organization has within its social network. The following two types of centrality are often used: point centrality and intermediate centrality. Point centrality can be measured if the focus is on interactive activities, while intermediate centrality can be measured if the focus is on control of interactions. To use social network analysis to identify opinion leaders, the most
important and commonly used concept is centrality. Therefore, for the analysis of centrality, this sub-section provides a detailed analysis of two different perspectives, point centrality, and intermediate centrality, as follows.

3.4.1 Degree centrality analysis

Degree centrality indicates the number of nodes in the network that are directly connected to the node, and the measure counts only the number of relationships that an actor or node has. The ability of a node to interact directly with other nodes and generate connections is usually expressed in terms of point degree centrality. If the size of a node's influence on network power is measured, it can be estimated by the point centrality metric.

If an actor is directly connected to a large number of community members, the higher the ability to interact directly and therefore the higher the point centrality, the actor corresponding to that point is a central figure and has a significant influence on other actors in the network. The actor corresponding to that point is marginal and has little influence on the community. Therefore, we usually measure the point centrality of actors to determine who qualifies as a key node in a network community.

<table>
<thead>
<tr>
<th>number</th>
<th>nodes</th>
<th>point centrality</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>People</td>
<td>15.000</td>
</tr>
<tr>
<td>2</td>
<td>CCTV</td>
<td>10.000</td>
</tr>
<tr>
<td>3</td>
<td>KNEWS</td>
<td>9.000</td>
</tr>
</tbody>
</table>

According to Table 2, the network "D2809 train incident" has the highest point centrality value of People(15.000), which means that this node has the most influence and is the node with the strongest information interaction and resource control. The next nodes are CCTV, KNEWS, ChinaNews, and zhiqun, whose point centrality is 10, 9, 8, and 7. Based on the results of point centrality measurement, it can be concluded that the above nodes are the opinion leaders of the "D2809 train incident" and have close and complex interactions with the nodes around them. They have a close and complex interaction with the nodes around them and have a great influence on the information dissemination and communication of the emergency online group.

3.4.2 Intermediate centrality analysis

Mesocentricity examines the extent to which an actor has control over information resources, i.e. the extent to which the actor has control over the capabilities of other actors, i.e. the actor's ability to intervene or control information interactions at other nodes. Intermediation is the extent to which the actor acts as a 'broker' for other actors. Those with a high degree of intermediation also have a high degree of intermediation centrality, acting as a bridge to other actors. Intermediation is therefore an indicator of 'control', which identifies which nodes in a social network can control the transmission of information. The more information links that need to be passed between other users via a particular user, the more central that user is to the network.
TABLE 3 RESULTS OF INTERMEDIATE CENTRALITY MEASUREMENTS FOR THE "D2809 TRAIN INCIDENT"

<table>
<thead>
<tr>
<th>number</th>
<th>nodes</th>
<th>Intermediate centrality</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CCTV</td>
<td>72.923</td>
</tr>
<tr>
<td>2</td>
<td>ChinaNews</td>
<td>49.591</td>
</tr>
<tr>
<td>3</td>
<td>Beijing</td>
<td>37.830</td>
</tr>
</tbody>
</table>

According to Table 3, we can learn that the nodes with the highest intermediate centrality are CCTV, ChinaNews, Beijing, People, and JIMU. These nodes are at the center of controlling the spread of public opinion and play the role of a bridge. In addition, several nodes in the network of the D2809 train incident have zero centrality, which indicates that they have little ability to control the relationships. They are at the periphery of the network, with relatively limited resources and limited access to information, and do not act as bridges of communication.

3.4.3 Proximity centrality analysis

Proximity to the center is a measure of how well a node in a network can propagate information without being 'controlled' by other nodes, and the smaller the value, the more centrally located the point is. The shorter the distance between a node and all other nodes, the higher the proximity of the node to the center. This is because a non-central node can be easily controlled by other nodes if it has to transmit information via other nodes. This means that a node that is connected to other nodes via a relatively short path can be identified as having a high degree of proximity centrality.

TABLE 4 RESULTS OF THE "D2809 TRAIN INCIDENT" PROXIMITY CENTRALITY MEASUREMENT

<table>
<thead>
<tr>
<th>number</th>
<th>nodes</th>
<th>Proximity centrality</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CCTV</td>
<td>36.000</td>
</tr>
<tr>
<td>2</td>
<td>ChinaNews</td>
<td>39.000</td>
</tr>
<tr>
<td>3</td>
<td>Beijing</td>
<td>39.000</td>
</tr>
</tbody>
</table>

According to Table 4(top3 in ascending order of centrality), the least influential and least controlled nodes in the "D2809 train incident" were the actors CCTV, ChinaNews, Beijing, People, and dianke. At the same time, it can be found that the proximity centrality of the top 10 actors, the point centrality of the top 10 actors, and the intermediate centrality of the top 10 actors have many similarities. This means that these nodes are located at the center of the opinion dissemination network and other actors cannot easily control them. On the contrary, they will have a significant influence on other actors in the dissemination of information, and they are important opinion leaders.

4 TIME SERIES ANALYSIS

In this paper, we analyze the trend of user sentiment evolution in the process of public opinion dissemination through time-series analysis, taking user sentiment change as the entry point to study the evolution of user sentiment at different points in time. The scatter diagram of the
sentiment distribution of microblog posts is shown in Figure 2. The sentiment score ranges from [0,1], where a value closer to 1 means the text content is more positive, and closer to 0 means the text content is more negative. Using the positive emotion score as a criterion, a positive emotion score in the range of [0,0.35] was judged as negative, a neutral emotion in the range of [0.35,0.65], and a positive emotion in the range of (0.65,1]. According to the calculation results, there were 1235 tweets with positive sentiment, accounting for 52.82% of the total number of tweets; 442 tweets with the neutral sentiment, accounting for 18.91% of the total number of tweets; and 661 tweets with negative sentiment, accounting for 28.27% of the total number of tweets.

Figure 3 shows the accumulation area of microblog posts within 11 days after the "D2809 train incident", from which we can visualize the "breaking point" of online users' emotions caused by the spread of online public opinion. From June 4 to June 14, the sentiment value of online users can be divided into three stages, namely "June 4-June 5 (1d)”, "June 5-June 9 (5d)”, "June 9-June 11 (3d)” and "June 9-June 11 (3d)”. June 9 - June 11 (3d)”. In the first phase, user sentiment was extremely positive, with topics such as "#D2809 accident rescue scene" and "#D2809#” focusing on the disaster accident being highly trendy during the period immediately after the accident; in the second phase, user sentiment was stable, with the topics "#d2809 train skidded more than 900 meters" and "#d2809 accident update summary" on the top search, and users showed higher concern for the details of the cause of the accident; in the third stage, sentiment fluctuated, and Most of the sentiments were positive. In particular, on June 10, 2022, the State General Administration of Railways announced the cause of the train dislodgement accident, and the announcement of these events triggered heated discussions among the general public and remembrance of the train driver, so users' sentiment fluctuated greatly, especially users' negative sentiment rose more. after June 11, 2022, people's sentiments tended to calm down and the heat of the event was also subsiding.
5 CONCLUSIONS

As a new type of information exchange and dissemination platform, the Internet has naturally formed some "opinion leaders". The Internet public is the main group of people involved in Internet activities and also the main group of people involved in public opinion on the Internet during public emergencies, and they play a huge role in influencing the spread of public opinion. During public emergencies, the netizen public often displays many irrational behaviors, usually showing that they are prone to believe rumors, are easily implied, and tend to think of simple things as complicated. The reason for these irrational behaviors is that, in the course of a public emergency, there is an imbalance in the netizens’ perceptions, and they do not believe that a crisis has arrived; at the same time, they are more likely to believe information that is favorable to them and to follow it blindly. The reasons for this are, on the one hand, the lack of authoritative and accurate information and, on the other hand, the lack of awareness of risk, responsibility, and the overall situation among the majority of Internet users. Based on the above research, we propose the following recommendations: First, give full play to the mainstream influence of the official media, which shoulders the important mission of guiding online public opinion on public emergencies. Secondly, the role of online opinion leaders should be given full play. Third, give full play to the positive guiding power of the self-media. The self-media must adhere to the correct political direction, put their central work in place, spread the voice of China with strong cohesion and leadership, promote the positive energy and main theme of society, and constantly improve their political discernment and ability to guide online public opinion.

![Weibo posting sentiment trend graph](image.png)

Figure 3. Weibo posting sentiment trend graph

REFERENCES

Cause Analysis of Ship Accidents Based on Chi-Square Test

Xintong Liu, Qing Yu, Yutian Hu, Wenting Lu, Aipan Zhang, Xujiang Bo*

* Corresponding author: xujianbo_0@hotmail.com

School of navigation, Jimei University, Xiamen, Fujian, China, 361000

Abstract: Studying historical accidents is of paramount significance to safeguard maritime transport. This study investigates the causation factors and their interactions from collected maritime accidents, which happened in the coastal water of China, to provide effective references for accident prevention. A chi-square test is used to assess the degree of association between two factors so that the potential causal relationships among major factors can be revealed. Then, the importance of each factor is ranked to identify the critical factors based on historical accident records. The research findings are expected to improve the understanding of the causation of the maritime accident while supporting maritime safety assessments and ship management.

Keywords: Chi-Square Test, Methods Use, Cause Analysis, Marine Accident.

1. INTRODUCTION

Nowadays, Waterborne transportation has been actively promoting the development of the national economy as an important mode of transportation, accounting for approximately 90% of the world’s trade in volume [1]. However, catastrophic accidents are frequently noted during the past years with more and more challenging safety issues. For instance, in the accident of Suwatai 999, the ship sunk and brought catastrophic economic loss [2]. Maritime accidents are commonly understood as the results caused by complex technical, human, organizational, and environmental factors [3]. Thereby, each severe maritime accident shall be analyzed to take experience from the accidents and to form a basis for the conventions and contracts produced for the prevention [4]. However, previous studies only discussed the statistical characteristics of the accidents [5], without a deeper study of the association between each variable. To study the causes of maritime accidents, this study applies a chi-square test to analyze the historical accident records from the coastal water of China. The chi-square test is a versatile hypothesis testing method. The relationship between the two variables will be analyzed by the chi-square test to detect the importance of each causation factor.

The remainder of this paper is structured as follows. The methodology is presented in Section 2; Section 3 provides a case study and the conclusion is drawn in Section 4.
2. METHODOLOGY

The chi-square test is used to analyze the interactions among all independent factors and the fitted distributions, which outputs the goodness of fit of a sample. The typical methods available for the chi-square tests include the Karl Pearson family, the Yates chi-square test, the Mantel–Haenszel chi-square, and the Maxwell–Stuart tests. One commonality of the above-mentioned methods is they use chi-square distribution as the reference distribution, which means that the chi-square test compares an observed set of data to what is expected [6].

Three parameters are used in the Karl Pearson family of chi-square tests: Goodness of fit, independence, and homogeneity. Each of these three tests has different interpretations and assumptions. The parameters can be calculated by using the following function:

$$x^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

where O is the target factor and E is the compared factor, n is the number of cells in the table, in which $i \in n$. The difference and usages among the three chi-square tests relate to the applicable situations and specific problems. The chi-square goodness of fit test is used to compare a sample to a population with known parameters on the variable of interest, or, to test the independence and to detect the correlation between two factors. In addition, homogeneity is used to detect distributional variability in two or more independent samples on individual variables of interest. Its common usage is to compare two or more groups on classification results.

3. CASE STUDY

A total of 9 causation factors are selected and compared to study the interactions and their importance among them. They are ship type, accident type, day or night, visibility, extreme weather conditions, seasons, weather, ship density, and human factors. The accident records are collected from public resources, such as MSA and Marine Bureau’s official website. Before processing the chi-square test, the obtained raw data need to be cleared by deleting the invalid data and errors. As result, a total of 108 records are used in this study.

The study selects the factor of "accident severity" as the target. Other factors of "ship type", "accident type", "day or night", "visibility", "extreme weather conditions", "seasons", "weather", "ship density", "human factors" and "count" are compared with the target factor.

Then states are assigned to all factors. The states for "ship type" include: 1) "general cargo ship", 2) "oil tanker", 3) "fishing vessel", 4) "Passenger ship", 5) "Bulk carrier", 6) "Other". For "accident type", states assigned are: 1) "Fire", 2) "Sinking", 3) "Grounding", 4) "Allision", 5) "Collision", 6) "Other". Assigned states for "Day or Night" as 1) "Day" and 2) "Night". Assign the states of "visibility" as 1) "bad", 2) "good" and 3) "medium". Assign states for "Extreme weather conditions" as 1) "None" and 2) "Yes". The factors of "seasons" as 1) "Spring", 2) "Summer", 3) "Autumn", 4) "Winter" respectively. The states for "Weather" are 1) "Sunny", 2)
"Cloudy", 3) "Fog", and 4) "Rain". The states for "ship density" are 1) "Low" and 2) "High". The states for "Human Factors" are 1) "No" and 2) "Yes".

After assigning states to each factor, this paper uses the SPSS software program to perform the chi-square test as well as a correlation analysis among these factors. A total of nine chi-square tests are performed.

If the chi-square results satisfy the result of P<0.05, it indicates the interaction between two factors. A smaller chi-square value indicates a smaller deviation between the two factors, while the opposite indicates a larger deviation. The results are shown in Table 1.

<table>
<thead>
<tr>
<th>Table 1 The chi-square value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factors</td>
</tr>
<tr>
<td>Day and night</td>
</tr>
<tr>
<td>The type of accident</td>
</tr>
<tr>
<td>Vessel type</td>
</tr>
<tr>
<td>Visibility</td>
</tr>
<tr>
<td>Extreme weather</td>
</tr>
<tr>
<td>Season</td>
</tr>
<tr>
<td>Weather</td>
</tr>
<tr>
<td>Density</td>
</tr>
<tr>
<td>Human factor</td>
</tr>
</tbody>
</table>

Table 1 reports the compared results of chi-square values among all factors. Three factors of day/night, type of ship accident, and ship type have significant interaction with "accident severity", with value less than the threshold value of 0.05 (0.017, 0.00003 and 0.027 respectively). The chi-square test results for other factors of visibility, extreme weather conditions, seasons, weather, ship density, and human factors are P=0.766, P=0.747, P=0.621, P=0.647, P=0.668, and P=0.747, respectively, in which the P-values of the above six factors are greater than 0.05, so there is no significant interaction between them and the severity of ship accidents.

In details, accident type, vessel type, and day or night are the three major influential factors of marine accidents in the coastal water of China, and the governments and the International Maritime Organization need to focus on these three factors and take effective measures. The frequency and proportions for all states of factors are reported in Table 2,3,4. There is the most significant correlation between the "types of accidents" and the "severity of the accident".

| Table 2 Type of accident as a percentage of accident severity |
|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| accident type       | Frequency | Proportions | Frequency | Proportions | Frequency | Proportions |
|                     | Level 1 | Level 2 | Level 3 | Level 1 | Level 2 | Level 3 |
| Fire                | Minor accidents | 4 | 0 | 0 | 4.80% | 0 | 0 |
|                     | Serious accidents | 0 | 4 | 5 | 40% | 40% | 35.70% |

Table 3 Type of vessel as a percentage of accident severity

<table>
<thead>
<tr>
<th>Vessel type</th>
<th>Frequency</th>
<th>Proportions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Level 1</td>
<td>Level 2</td>
</tr>
<tr>
<td></td>
<td>Minor</td>
<td>Serious</td>
</tr>
<tr>
<td>Grocery ships</td>
<td>18</td>
<td>0</td>
</tr>
<tr>
<td>Tanker</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Fishing vessel</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>Passenger ship</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Bulk carriers</td>
<td>31</td>
<td>2</td>
</tr>
<tr>
<td>Other</td>
<td>22</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 4 Day/night as a percentage of accident severity

<table>
<thead>
<tr>
<th>Day/night</th>
<th>Frequency</th>
<th>Proportions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Level 1</td>
<td>Level 2</td>
</tr>
<tr>
<td></td>
<td>Minor</td>
<td>Serious</td>
</tr>
<tr>
<td>Day</td>
<td>28</td>
<td>7</td>
</tr>
<tr>
<td>Night</td>
<td>56</td>
<td>3</td>
</tr>
</tbody>
</table>

From the above three tables, the important factor influencing accident severity and their inner relationships can be seen. The result shows collision is the major accident type that occurs in
coastal waters mainly happening on bulk carriers at night. Different state combinations of these three factors would lead to different accident severity. IMO should strengthen the cooperation among departments and accelerate the speed of salvage. Meanwhile, the ship safety inspection should be carried out to ensure that the ships are in condition, and enhanced communication between ships to reduce the occurrence of ship collisions. According to the data in this paper, the probability of accidents at night is higher than during the day because the crew is more likely to be tired at night, so the crew needs to maintain regular observation at all times of the night. Once a maritime accident is unavoidable relevant departments and rescuers should establish an effective plan to avoid secondary hazards.

4. CONCLUSION

This paper reports a study of causation analysis of ship accidents based on the chi-square test. The chi-square test is used to analyze the correlation between accident severity and other nine factors. As results, the day/night, type of accidents and ship type are three important factors. In addition, the states of those three factors are also analyzed. The test reveals some potential causal relations including collision is the major accident type that occurs in maritime bulk carriers; minor accidents often happen in bulk carriers. Catastrophic accidents usually happened at night. The study results are useful to provide effective references for accident prevention and to ensure safety sailing at sea. To the result would be more accurate, the future research database should be expanded to involve more data sources to increase the confidence level of experimental data.

REFERENCES

Analysis of Southeast Asian Potential Ports in the Container Shipping Network Between China and Southeast Asian Nations

Siheng Chang1, Jinhai Chen1, *, Peng Peng2

*Corresponding author: jhchen@jmu.edu.cn

Navigation College of Jimei University, Xiamen 361021, Fujian, China1
State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Beijing 100101, China2

Abstract: China and the Association of Southeast Asia Nations (ASEAN) are geographically nearby with a high degree of economic interdependence. Most of the Southeast Asian countries are emerging economies with a large demand for maritime infrastructures. Therefore, improving ports infrastructure in this region could substantially increase the growth of transportation and economic activities between China and ASEAN. This paper adopts the centrality index parameters using the characteristics of the port of the China-ASEAN shipping network and a reasonable calculation method of system centrality is put forward using actual ports throughput data. The 92 container ports in the China-ASEAN shipping network are graded based on the comprehensive evaluation and potential container ports that have been identified in Southeast Asia. The main findings are as follows:(1) Singapore Port and Kaohsiung Port are the hub ports of the China-ASEAN container shipping network, which control the development of the maritime trade between China and ASEAN;(2) Cai Mep Port in Vietnam owns the largest potential, as compared with other considered potential ports connected Chinese ports in Southeast Asia;(3) The 12 potential container ports in Southeast Asia mainly located in Vietnam and Philippines. For the Chinese government and enterprises, investing in port construction in Vietnam and the Philippines could be a smart choice.

Keywords: China-ASEAN Shipping Network, Ports, Maritime Big Data.

1 INTRODUCTION

With an initiative like the ‘21st-Century Maritime Silk Road (MSR) and the Regional Comprehensive Economic Partnership (RCEP) [1], China and ASEAN have built a solid and robust partnership with booming trade and economic cooperation, which has sped up infrastructure construction and improved interconnectivity over the last two decades. China and ASEAN are the world’s main economies, and they are also among the world's fastest-growing. In August 2020, ASEAN Secretariat notes its concatenated gross domestic product (GDP) of US $3.166 million in 2019 grew by 28.2% from the 2015 figure that stood at the US $2.469 million. It means the Southeast Asian region has huge development potential.

Maritime transportation is the major conduit of international trade, and ports play a key part in international cargo transportation [2]. Southeast Asia is a vital region of the global maritime
industry, particularly for cargo shipments between Asia and Europe. The Coronavirus disease (COVID-19) triggered a global health and economic disaster with far-reaching implications for international trade. However, 2020 was a banner year for China-ASEAN trade. ASEAN replaced the European Union as China’s top merchandise trading partner in 2020, according to statistics from the General Administration of Customs of China. Imports and exports between China and ASEAN will amount to 4.74 trillion yuan in 2020, increasing 7% year on year. Therefore, the Chinese government and businesses may greatly boost the growth of economic activities between China and ASEAN by increasing port investments in the Southeast Asian area. However, ASEAN has its internal issues, such as uneven economic progress among its members. Most Southeast Asian countries are developing countries, and port infrastructure construction is facing various challenges notably a shortage of funds. This raises a significant and pressing problem of identifying Southeast Asia’s potential ports connected with Chinese ports. It is critical to increasing investments and infrastructure construction in these ports to maintain the economic viability of China’s connections with Southeast Asian countries.

The application of the centrality concept in maritime container transportation, in particular to the port sector, has already been addressed in the literature. Fleming and Hayuth first introduced port centrality and intermediacy as locational attributes indicative of the strategic role of each port within a transportation system [3]. Centrality was defined as the locational advantage within the market area the port serves. In other words, if the port is located in the midst, rather than being on the periphery, of a large hinterland, then the more central port has an advantage in attracting extra traffic generated from that hinterland. Intermediacy represents the natural geographical ‘in betweenness’ of a port in connecting more than one foreland market. Intermediate ports can attract extra traffic if they are favored by carriers as connecting hubs or relay points in the system. Typical examples are the ports of Hong Kong or Singapore, which are strategically located in favorable intermediate positions along major sea trade routes. The concept of port centrality has greatly facilitated practical research related to port development and inter-port relationship. For example, the feature of port betweenness centrality and degree centrality have been utilized by Hu and Zhu to identify the potentially congested ports in a busy maritime network [4]. Some studies provide further evidence of the effectiveness of port centrality in determining a port hierarchy and in indicating all the attributes underpinning it [5-8]. A major shortcoming of port centrality studies is there is little comprehensive quantitative benchmarking available for assessment. This study aims to adapt a comprehensive grade method based on complex network theory and centrality index and take port throughput into account. We identify potential container ports connect with Chinese ports in Southeast Asia.

2 METHODOLOGY AND DATA

2.1 Network Construction

This study focuses on the degree of trade relations between ports; hence, we employed a directed graph, $G = (V, E)$ that represents the overall set of between China-ASEAN shipping network trade relations constructed on an annual time scale. For the graph $G = (V, E)$, $V=\{v_1, v_2, \cdots v_n\}$ is set of all ports in $G$, and is the set of all edges that link the pairs of ports in $V$, representing the routes between ports $i$ and $j$. The elements in the adjacency matrix take the form of $e_{ij}=1$; if a route exists; otherwise, $e_{ij}=0$. 


2.2 Centrality Indicators

(1) Degree Centrality
The degree of centrality is a measure of how many of the nodes in a network are connected. The connection centrality of a node refers to network connectivity. A port’s degree centrality refers to the number of neighbor ports directly connected with the port. The degree centrality can be represented as the “organizational capacity” of a container port. If node $i$ and node $j$ are connected, $a_{ij} = 1$ can be defined; if not, $a_{ij} = 0$ can be defined as follow.

$$DC_i = \frac{1}{n-1} \sum_{j=1,j\neq i}^{n} a_{ij}$$

(2) Betweenness centrality
Betweenness centrality is defined as the proportion of the shortest paths between every pair of nodes that pass through the given network towards all the shortest paths. It mainly reflects the influence of nodes in the entire network, as high values often correspond to hubs or bridges.

$$BC_i = \sum_{s \neq i \neq t} \frac{\sigma_{st}(i)}{\sigma_{st}}$$

where $\sigma_{st}(i)$ is the sum of the shortest paths from nodes $s$ to $t$ that pass through node $i$ and $\sigma_{st}$ is the sum of the shortest paths from nodes $s$ to $t$.

(3) Closeness centrality
Closeness centrality measures the minimum distance between a given node and other nodes, which reflects the relative accessibility of that node in the network. The closeness centrality can evaluate the role a node plays in a network. A container port is connected with other container ports by the shipping lines, so the closeness centrality not only means the shipping lines coverage of a container port but also reflects the “shipping accessibility” of a container port and its significance in the global container shipping network. The greater the closeness centrality of the port, the higher the relative reachability of the port.

$$CC_i = \frac{1}{n-1} \sum_{j=1,j\neq i}^{n} d_{ij}$$

where $d_{ij}$ is the length of the topologically shortest path between ports $i$ and $j$.

(4) Eigenvector Centrality
Eigenvector centrality emphasizes the importance of a network node. It measures the relative score of all the nodes in the network based on the principle that the connection to a node with a higher centrality value has a greater effect on the centrality score. High eigenvector centrality
means that, in a network, the central port not only has a large number of connecting routes, its connecting ports also have a significant impact on it.

\[
EC_i = \frac{1}{2} \sum_{j=1}^{N} a_{ij} x_j
\]  

(4)

where \( \lambda \) is the constant, \( a_{ij} \) is one if \( j \) is connected to \( i \) and zero otherwise. \( x_j \) is the centrality of node \( j \).

(5) System Centrality

Centrality is one of the most studied concepts in complex network research\([9][10]\). Centrality measures might be generally classified into four basic categories relating to different perspectives: degree centrality, closeness centrality, betweenness centrality, and eigenvector centrality, which analyze transportation networks to evaluate the status of nodes in the network as a whole or at the level of nodes\([11][12]\). However, the centrality indicators of degree, closeness, betweenness, and eigenvector represent a node’s location advantage as being directly connected to others, being accessible to others, being the intermediary between others, and the importance of ports in terms of its connectivity with important ports, respectively. Any central indicator can only reflect topological features, but not comprehensive, to comprehensively reflect the hub ability of the port, this paper draws lessons from Han et al\([13]\). They put a method that the System Centrality (\( SC_i \)) on the study stations classification of China’s high-speed network, which is the comprehensive quantitative centrality.

The SC is calculated as follows:

\[
SC_i = (\omega_1 \cdot CC_i + \omega_2 BC_i) \times DC_i
\]  

(5)

where \( CC_i, BC_i, DC_i \) is closeness centrality, betweenness centrality, degree centrality, respectively. \( \omega_1, \omega_2 \) is the weighted coefficient of \( CC_i, BC_i \).

According to the characteristics of the China-ASEAN shipping network and ports’ throughput, we consider eigenvector centrality indicators. The new parameters are correspondingly adapted by giving due consideration to the container port’s throughput that is only involved between China and Southeast Asia.

We are making correlation analysis for degree centrality, closeness, centrality, betweenness centrality, eigenvector centrality with container port’s throughput. The results are shown in Table1. There exists the highest correlation (0.884) between the betweenness centralities and container port’s throughput of container ports and the lowest correlation (0.379) between the closeness centralities and container port’s throughput of container ports, which indicates that the characteristics of container cargo transport, with traffic size and reachability, are not the primary objectives of maritime cargo transport, and high transshipment is quite attractive for the organizational effectiveness of ports. We also consider the polarization of betweenness centrality, which is a node with better betweenness centrality in the network that will have its system centrality enlarged (e.g. multiplier effect). So the final formula is
\[ SC_i = (\omega_1 CC_i + \omega_2 DC_i + \omega_3 EC_i) \times (1 + BC_i) \]  \hspace{1cm} (6)

Where \( \omega_1 \), \( \omega_2 \), \( \omega_3 \) is \( CC_i \), \( DC_i \), \( EC_i \) weighted coefficient. \( \omega_1=0.1, \omega_2=0.45, \omega_3=0.45 \).

We carry out a correlation analysis between the result of SC and corresponding port throughput. The Pearson correlation coefficient is 0.750, which indicates that system centrality is not only relevant to the role of the port in the network but also has a strong correlation with the port throughput.

<table>
<thead>
<tr>
<th>Correlation coefficient(R)</th>
<th>DC</th>
<th>EC</th>
<th>BC</th>
<th>CC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Throughput</td>
<td>0.790</td>
<td>0.666</td>
<td>0.884</td>
<td>0.379</td>
</tr>
</tbody>
</table>

### 2.3 Data

This paper analyzed the China-ASEAN shipping network, involving 93 container ports and 9193 OD routes in 2019 (www.shipxy.com). We used these data to calculate the arrival and departure records of vessels at all calling container ports. This resulted in a dataset comprising all ports, routes, and journeys, which was used to define the transportation networks.

### 3 RESULTS

The container ports mainly located in eastern coast of China, Indonesia, Philippines, Vietnam and Thailand in the China-ASEAN shipping network (in Fig.1).

![Image of distribution of ports in China-ASEAN shipping network](image-url)
Table 2 The hierarchical structure of the ports system in China and Southeast Asia

<table>
<thead>
<tr>
<th>Level</th>
<th>Scores</th>
<th>Number of Ports</th>
<th>Ports</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>[0.8,1]</td>
<td>2</td>
<td>Singapore; Kaohsiung; Manila; Ningbo; Shanghai; Xiamen;</td>
</tr>
<tr>
<td>2</td>
<td>(0.45,0.8)</td>
<td>14</td>
<td>Tanjung Pelepas; Hong Kong; Qingdao; Cat Lai; Laem Chabang; Tanjung Priok; Port Klang; Haiphong; etc.</td>
</tr>
<tr>
<td>3</td>
<td>(0.15,0.45)</td>
<td>28</td>
<td>Cai Mep; Saigon; Taichung; Subic Bay; Batangas; Zhou Shan; Danang; Vung Tau; Mawan; Fuzhou; Bintulu; etc.; Quanzhou; Shantou; Shenzhen;</td>
</tr>
<tr>
<td>4</td>
<td>(0,0.15)</td>
<td>49</td>
<td>Davao; Samarinda; Rizhao; Lanshan; Kuching; Keelung; etc.</td>
</tr>
</tbody>
</table>

Table 2 shows all container ports in China and Southeast Asia are classified into four layers, focusing on the first to third layers. The first layers as the container ports hub ports, the second layers as regional hub ports, and the third layers as the ports with development potential in the China-ASEAN container shipping network. The result shows that Singapore and Kaohsiung are the hub ports of the China-ASEAN container shipping network and control the development of the maritime trade between China and ASEAN. Hub ports in the first layer and regional hub ports the second layer (Fig.4) are located in China (Kaohsiung Port, Ningbo Port, Shanghai Port, Xiamen Port, Hong Kong Port, Qingdao Port, Shekou Port Guangzhou Port, Singapore (Singapore Port), Philippines (Manila Port), Malaysia (Tanjung Pelepas Port, Port Klang,), Vietnam (Cat Lai Port, Haiphong Port), Thailand (Laem Chabang Port), Indonesia (Tanjung Priok Port). Port Klang, Tanjung Pelepas, Laem Chabang are emerging as new competitors to Singapore and could become new hub ports. Specifically, the Singapore Port, Port Klang and Tanjung Pelepas Port of Malaysia in the Strait of Malacca are the primary and secondary chokepoints in both core and secondary maritime shipping routes in this region. These are the top container ports in Southeast Asia, all with well-developed infrastructure, such as Port Klang and Tanjung Pelepas port upgraded their existing facilities, infrastructure and scale by building new berths and container yards to handle over 20.8 million TEU in 2018.

In the first and second layers (in Fig.2), there is a high eigenvector centrality of each container port in the China-ASEAN shipping network. It is observed that Singapore has the highest eigenvector centrality. Specifically, the ports of eigenvector centrality in the first and second layers much higher than others. Therefore, it demonstrates that most container ports have limited ability to open up direct shipping lines to key container ports and conversely most shipping trunk lines concentrate on a few container ports in the China-ASEAN shipping network. Singapore (0.172) has the highest betweenness centrality. The container port with the second and third highest betweenness centrality is Kaohsiung (0.141), Manila (0.08) respectively. The betweenness centrality of Manila has a big gap compared to Singapore Port and Kaohsiung Port. Undoubtedly, Singapore Port and Kaohsiung Port have become the transshipment hub port for China and the Southeast Asia countries.

In the third layer (in Fig.3), we focus on 12 potential container ports connected with Chinese ports in Southeast Asia, which distribute in Vietnam (Cai Mep Port, Saigon Port, Danang Port, Vung Tau Port), Philippines (Subic Bay Port, Batangas Port, Cebu Port, Cagayan de Oro Port), Thailand (Bangkok Port, Phuket Port), Malaysia (Johor Port, Bintulu Port), Indonesia (Tanjung
Perak Port). These ports with higher eigenvector centrality, as compared with fourth lawyer ports, which means that the ports not only have some connecting routes, its connecting ports also have a significant impact on it. Higher degree centrality also reflects the potential ports that can connect the direct shipping lines to the other neighbor ports. Variation of closeness centrality smoothly indicates these potential ports are mainly located near the shortest paths of many Origin-Destination pairs between China and Southeast Asia. Furthermore, one can find that Vietnam and the Philippines have the most ports with development potential. The main

**Fig. 2** Ports in the first and second layers

**Fig. 3** Ports in third layers located in Southeast Asia
reason that the digital economy spawned by the pandemic in 2020, cooperating more tightly in the electronics manufacturing industry between China and ASEAN. Especially, electrical and mechanical products are the first major product in bilateral trade between China and Vietnam. In 2020, Vietnam also serves as China’s largest trading partner in the ASEAN. Vietnam is one of the most important suppliers of rice to China; China was still the Philippines’ largest trading partner in the first (fourth-largest export market, largest import source) and second quarter (second-largest export market and largest import source) in 2020. Both countries have also agreed to establish a “fast track” lane and a “green channel” to guarantee the steady flow of goods and stability of industry a land supply chains between the Philippines and China impacted by COVID-19. Thus, benefited from both economic and geographic proximity.

The Cai Mep Port in Vietnam has the largest development potential, as compared with other considered potential ports. Take a closer look, the eigenvector centrality of Cai Mep Port is 0.573 and the degree centrality is 0.44. It indicates Cai Mep Port connects a certain number of hub ports and regional hub ports and relatively closes trade links in the China-ASEAN shipping network. Cai Mep Port is strategically located 50 kilometers southeast of Vietnam’s commercial hub, Ho Chi Minh City. It offers direct access to and from Asia, Europe, and the Americas. These advantages prove it has large development potential and value investment.

In the fourth layer(in Fig.4), the top 10 Southeast Asia ports in the fourth layers mainly distribute in Indonesia(Samarinda Port, Makassar Port, Balikpapan Port, Cigading Port, Muara Pantai Port ). Why Indonesia ‘s ports lack advantages compare with Vietnam and the Philippines in the China-ASEAN shipping network?

![Fig.4 Top 10 ports in the fourth layers located in Southeast Asia](image)

During the pandemic, Indonesia and China’s economic relationship is seen to keep progressing despite both countries’ difficulties in a domestic economic context. Fortunately, in commodities, the trade between Indonesia and China is still growing. In detail, 11 Indonesian commodities showed a significant increase in the first quarter of 2020 as shown by Table 3. As shown in Table 3, Packed fruits grow fastest, but electronics slowest. The Ministry Of Commerce People’s Republic of China notes, In recent years, China and ASEAN keep progressing on different
industrial chains, particularly in the electronics manufacturing industry, which has a significant growth in the value of associated product imports and exports. As a part of the global electronics manufacturing chain, China imported 226.81 billion yuan of integrated circuits from ASEAN in the first half of 2020, increasing 23.8 percent year on year and accounting for 24.2 percent of overall ASEAN imports. The integrated circuit, energy, agricultural products contributed 3.2 %, 1.7 %, and 0.8% to China's trade growth with ASEAN, respectively. Thus, the small percentage of export growth electronics is one of the reasons for the lack of advantages of its ports.

Table 3 Indonesian commodity export rise

<table>
<thead>
<tr>
<th>No.</th>
<th>Commodity</th>
<th>Increase percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Packed fruits</td>
<td>320.27</td>
</tr>
<tr>
<td>2</td>
<td>Processed woods</td>
<td>222.44</td>
</tr>
<tr>
<td>3</td>
<td>Iron and steel</td>
<td>196.40</td>
</tr>
<tr>
<td>4</td>
<td>Swift’s nest</td>
<td>189.61</td>
</tr>
<tr>
<td>5</td>
<td>Canned fish</td>
<td>92.59</td>
</tr>
<tr>
<td>6</td>
<td>Coal</td>
<td>74.42</td>
</tr>
<tr>
<td>7</td>
<td>Frozen fish</td>
<td>53.78</td>
</tr>
<tr>
<td>8</td>
<td>Shoes</td>
<td>24.59</td>
</tr>
<tr>
<td>9</td>
<td>Furniture</td>
<td>30.87</td>
</tr>
<tr>
<td>10</td>
<td>Tropical food</td>
<td>22.29</td>
</tr>
<tr>
<td>11</td>
<td>Electronics</td>
<td>14.70</td>
</tr>
</tbody>
</table>

One finding is that there is no great difference in the reachability of each container port which means the probability that one container port is connected with other container ports. The closeness centrality of China and ASEAN countries can be better understood if considering its geographic location, which indicates that transportation between almost all pairs of ports can be achieved directly or depending on one other intermediary ports and high reachability.

According to our results, we can further have the following suggestions:

(1) The national cooperation between Vietnam and China, the Philippines, and China have huge development potential and should be significantly improved shortly. As compared with other Southeast Asian countries,

(2) Investing in port construction in Vietnam and the Philippines can be a good choice for China to accelerate the implementation of the MSR and RCEP initiative.

4 CONCLUSION

Ports that serve as a connection point between the sea and hinterland transport constitute strong support for both logistical and economic activities. With the strengthening of trade and regional integration dynamics between China and Southeast Asian countries, direct contact has increased. Southeast Asia gradually became the most important investment area for the Chinese government and enterprise. This paper pays more attention to container ports connected with Chinese ports in Southeast Asia based on centrality analysis. The conclusion demonstrates that:

(1) Singapore and Kaohsiung are the hub ports of the China-ASEAN container shipping network and control the development of the maritime trade between China and ASEAN.
(2) Regional Hub ports are mainly located in China, the Philippines, Malaysia, Vietnam, Thailand, Indonesia. Potential ports are mainly located in Vietnam and the Philippines, Cai Mep in Vietnam has the largest investment potential, as compared with other considered potential ports connected to Chinese ports in Southeast Asia.

(3) As compared with other Southeast Asian countries, the national cooperation between Vietnam and China, the Philippines, and China have huge development potential in the future for the Chinese government and enterprise, investing port construction in Vietnam and the Philippines can be a good choice for China to accelerate the implementation of the MSR and RCEP initiative.

Acknowledgements: This research was supported by the Natural Science Foundation of Fujian Province (2020J01659), National Natural Science Foundation of China (42001391).

REFERENCES

A Visual Analysis of Chinese Internet Healthcare Research Based on Bibliometrics

Yufang He¹,a, Kaiyue Shen¹,b, Chengye Zhang¹,c, Hongjuan Wen¹,d*

1134707402@qq.com, Corresponding author: 1007481182@qq.com
d*

School of Health Management, Changchun University of Chinese Medicine, Changchun, China¹

Abstract: This study aims to explore the current status, hotspots and trends of research in the field of Internet healthcare in China. We analyzed 1796 related literatures in CNKI database through COOC and Vosviewer visualization tools. The results show that the number of literature is growing and can be roughly divided into three phases: germination, development, and new venture development; the core researchers are generally scattered and national inter-institutional cooperation has not been formed yet; the research hotspots mainly focus on three themes of internet medical information construction, platform application construction and service model management. Further research might focus on digital hospital construction and exploring demand-oriented medical services.

Keywords: Internet Healthcare, Visualization Research, COOC, Literature Metrics.

1 INTRODUCTION

Internet healthcare is a general term for a new medical and health service model formed by the deep integration of the Internet as a carrier and information technology with traditional medical services [1]. It plays an important role in improving the medical experience and solving the problems of patients' difficulties in accessing medical care, especially in the ongoing prevention and control of COVID-19. Internet healthcare has shown its unique advantages in maximizing the use of medical resources, reducing crowd gathering and avoiding cross-infection [2]. As Internet technology continues to be integrated with the medical field, more research has been conducted in the field of Internet healthcare, including the analysis of the construction path of Internet hospital informatization [3] and the exploration of Internet-based methods for chronic disease health management [4], but there are few macroscopic reviews of the development path of this field in China. In this paper, with the help of the COOC visualization tool, we will review the general situation of the Internet healthcare field in China from 2011 to 2021, discuss the hotspots and trends of research, to provide a reference for further research in this area.
2 MATERIALS AND METHODS

The data studied in this paper come from CNKI Chinese database, searched with the theme of "Internet medical care", the time threshold is limited to 2011-2021, the literature source category is selected as all journals, by removing documents with missing information (authors, keywords, institutions, etc.) and those not related to the topic, we ended up with 1796 valid documents. Valid literature was exported in Refworks format to form a sample database in the field of Internet healthcare research. Synonym merging, frequency statistics, co-occurrence matrices, clustering plots and time zone plots were performed using COOC 12.6 [5] and Vosviewer software to draw visual plots for visual analysis.

3 RESULTS AND DISCUSSION

3.1 Annual Publication Volume

From Fig.1, it can be seen that the research in the domestic Internet medical field has an overall growth trend. The period from 2011 to 2014 was a budding phase with a low literature output. During the period from 2015 to 2019, the number of articles begins to grow, the Internet medical market gradually expanded and gradually received the attention of many scholars. Since 2020, the "Internet+ Health" has been accelerated due to the impact of COVID-19, more and more scholars have started to pay close attention to the field of Internet healthcare, and the number of publications has increased.

3.2 Literature Source Journals

Table I shows, the nascent core publication position in the field of Internet healthcare in China has basically taken shape, with 1,796 Internet healthcare literature published in 629 journals including medical, health management and computer journals, "China Digital Medicine" contains the most relevant literature with a total of 75 articles, followed by 44 articles in "China Hospital".
### Table I. Distribution of the Top 10 Publishing Journals

<table>
<thead>
<tr>
<th>No.</th>
<th>Journal Distribution</th>
<th>Amount</th>
<th>Kind</th>
<th>Scale (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>China Digital Medicine</td>
<td>75</td>
<td></td>
<td>4.18</td>
</tr>
<tr>
<td>2</td>
<td>China Hospital</td>
<td>44</td>
<td>Core Journals</td>
<td>2.45</td>
</tr>
<tr>
<td>3</td>
<td>Chinese Health Information</td>
<td>43</td>
<td></td>
<td>2.39</td>
</tr>
<tr>
<td></td>
<td>Management</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Journal of Medical Informatics</td>
<td>38</td>
<td></td>
<td>2.12</td>
</tr>
<tr>
<td>5</td>
<td>Chinese Journal of Hospital</td>
<td>29</td>
<td>Core Journals</td>
<td>1.61</td>
</tr>
<tr>
<td></td>
<td>Management</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>China Health</td>
<td>27</td>
<td></td>
<td>1.50</td>
</tr>
<tr>
<td>7</td>
<td>China Health Industry</td>
<td>26</td>
<td></td>
<td>1.45</td>
</tr>
<tr>
<td>8</td>
<td>China Health Quality Management</td>
<td>23</td>
<td></td>
<td>1.28</td>
</tr>
<tr>
<td>9</td>
<td>Hospital Management in China</td>
<td>22</td>
<td>Core Journals</td>
<td>1.22</td>
</tr>
<tr>
<td>10</td>
<td>Information and Computers</td>
<td>21</td>
<td></td>
<td>1.17</td>
</tr>
</tbody>
</table>

**Total:** 348 19.38

### 3.3 Author Network Analysis

A total of 4015 authors were included in the valid literature, and according to Price's theory, three or more publications are core authors in the field, with a total of 133 core authors for whom a co-occurrence diagram of researchers with collaborative relationships was constructed (Fig.2). The co-occurrence analysis shows that in the field of Internet healthcare, a research team has been formed with Yu Guangjun (19), Liu Lifei (3), Zhang Xinping (3), Lu Wei (6), Miao Wei (7) and Yu Junying (4) as the core, while there are also scholars with a large number of publications, such as Cui Wenbin (12), Zheng Xueqian (9), Liu Yang (8) and Chen Min (8).

**Figure 2. Author Co-occurrence Map**
### 3.4 Institutional Network Analysis

Table II shows the top five organizations in terms of the number of articles published, all of which have more than ten articles, constituting the main position of research in the field of Internet healthcare. The 122 institutions with three or more publications were selected to build a collaborative network map (Fig.3), which revealed a team of research institutions with the School of Public Health of Shanghai Jiao Tong University, the First Hospital of Sun Yat-sen University and Sun Yat-sen Hospital of Sun Yat-sen University as the core.

**TABLE II.  **TOP5 RESEARCH INSTITUTIONS

<table>
<thead>
<tr>
<th>No.</th>
<th>Institution</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>School of Medicine and Health Management, Tongji Medical College, Huazhong University of Science and Technology</td>
<td>23</td>
</tr>
<tr>
<td>2</td>
<td>School of Health Management and Education, Capital University of Medical Sciences</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>School of Management, Shanghai University of Engineering and Technology</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>School of Public Health, Shanghai Jiao Tong University</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>Institute of Medical Information, Chinese Academy of Medical Sciences</td>
<td>10</td>
</tr>
</tbody>
</table>

![Figure 3. Institutional Cooperation Diagram](image)

### 3.5 High-Cited Literature

Table III lists the top ten cited articles that have made significant contributions to Internet health research. The research content of the highly cited literature can be divided into three categories: The construction of application service model, Meng Qun [6-7] pointed out that platform construction in areas such as graded diagnosis and treatment and chronic disease management would promote medical development, Chen Huifang [8] took registration as an example to carry out innovative practice of service model, and He Xuesong [9] analyzed medical core business and auxiliary business applications. Industry development trend, Xu Zhirong [10] divided the Internet medical market pattern into four categories, Guo Wei [11] explained the...
realistic positioning of Internet medical, Xie Wenzhao \cite{12} pointed out four construction modes and application modes of Internet medical in China, and Yu Baorong \cite{13} explored the operation form of Internet medical enterprises. The construction of the regulatory system, Wang Anqi \cite{14} and Li Ying \cite{15} made substantive suggestions on the supporting policies and related legislation in the field of Internet healthcare.

**TABLE III. TOP 10 HIGH-CITED LITERATURE ON INTERNET HEALTHCARE**

<table>
<thead>
<tr>
<th>No.</th>
<th>Topic</th>
<th>Author</th>
<th>Year</th>
<th>Journal</th>
<th>Citations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Analysis of innovation and opportunities for traditional industries in the era of &quot;Internet+&quot;</td>
<td>Xu Jianrong</td>
<td>2015</td>
<td>Internet World</td>
<td>244</td>
</tr>
<tr>
<td>2</td>
<td>Current situation and thoughts on the development of Internet healthcare in China</td>
<td>Meng Qun et al.</td>
<td>2016</td>
<td>Chinese Journal of Health Information Management</td>
<td>123</td>
</tr>
<tr>
<td>3</td>
<td>The current situation of Internet healthcare operation in China - a survey and analysis based on 3 hospitals</td>
<td>Wang Anqi et al.</td>
<td>2016</td>
<td>China Health Policy Study</td>
<td>113</td>
</tr>
<tr>
<td>4</td>
<td>Innovative development of &quot;Internet + Medical&quot;</td>
<td>Li Ying et al.</td>
<td>2016</td>
<td>Macroeconomic management</td>
<td>98</td>
</tr>
<tr>
<td>5</td>
<td>Realistic positioning and future development of Internet healthcare</td>
<td>Guo Wei et al.</td>
<td>2016</td>
<td>Explore</td>
<td>68</td>
</tr>
<tr>
<td>6</td>
<td>Research and Practice of Internet + Chronic Disease Management</td>
<td>Meng Qun et al.</td>
<td>2016</td>
<td>Chinese Journal of Health Information Management</td>
<td>66</td>
</tr>
<tr>
<td>7</td>
<td>The development status of Internet healthcare in China and the challenges it faces</td>
<td>Xie Wenzhao et al.</td>
<td>2016</td>
<td>Chinese Journal of Medical Library Information</td>
<td>64</td>
</tr>
<tr>
<td>8</td>
<td>The development history, business model and macro-influencing factors of Internet medical care in China</td>
<td>Yu Baorong et al.</td>
<td>2019</td>
<td>Journal of Shandong University (Medical Edition)</td>
<td>50</td>
</tr>
<tr>
<td>9</td>
<td>Research on Internet healthcare service models from the perspective of value co-creation</td>
<td>Chen Huifang et al.</td>
<td>2016</td>
<td>Modern management science</td>
<td>48</td>
</tr>
<tr>
<td>10</td>
<td>Current status and trends in the application of Internet healthcare</td>
<td>He Xuesong et al.</td>
<td>2018</td>
<td>China Health Policy Study</td>
<td>45</td>
</tr>
</tbody>
</table>
3.6 Keyword Analysis

Fig. 4 shows the top 20 keywords by frequency, the high-frequency keywords included Internet+, Internet medical, Internet hospital, mobile medical, smart medical, medical service, etc. These keywords are important hot content in the field of Internet medical research. A keyword co-occurrence map with a frequency of 5 times and above was constructed (Fig. 5), and there were 170 nodes and 1,186 connections in the map. Cluster analysis provides a fuller understanding of the thematic links between individual research hotspots, Table IV shows the five keyword clusters.

Figure 4. Statistics of the top 20 keywords in terms of word frequency

Figure 5. Keyword co-occurrence map
TABLE IV.  KEYWORD CO-OCCURRENCE NETWORK CLUSTERING TABLE

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Cluster size</th>
<th>Part of frequent keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>53</td>
<td>Graded treatment, medical platform, doctor-patient relationship, hospital information system, medical information security, chronic disease management, medical service system</td>
</tr>
<tr>
<td>1</td>
<td>45</td>
<td>Health care, cloud platform, regional medicine, precision medicine, community medicine, rehabilitation medicine, network medicine</td>
</tr>
<tr>
<td>2</td>
<td>33</td>
<td>Internet hospitals, medical and health services, artificial intelligence, palm hospitals, dingxiang garden, epidemic prevention and control</td>
</tr>
<tr>
<td>3</td>
<td>30</td>
<td>Mobile Health, telemedicine, healthcare information, outpatient services</td>
</tr>
<tr>
<td>4</td>
<td>9</td>
<td>Medical information, data security</td>
</tr>
</tbody>
</table>

Through the combing of relevant literature and clustering keywords, it is concluded that the current domestic Internet medical research mainly revolves around the following three themes:

Internet medical information technology construction: the common present symbols are "hospital information system", "medical information security", "medical data", "medical health information", etc. Insufficient total medical resources and unbalanced structural layout have always been the pain and difficulty of China's health care reform [16]. In recent years, hospitals at all levels in China are exploring information technology and intelligent medicine, especially during the prevention and control of COVID-19, the research of medical information technology construction has developed rapidly, and the construction of medical information technology effectively expands the diagnosis and treatment information outwardly and connects inwardly, which is an important link in the development of Internet medical and health care. However, the research on the security regulation of medical information data still needs to be improved, a specialized legal regulatory system has not yet been formed, and the legal rank is low.

Internet medical platform construction: the common symbolic words are "medical platform", "palm hospital", "cloud platform", "Dingxiangyuan", etc. There are two types of Internet hospital platforms, namely hospital-based and enterprise-based. In recent years, various forms of medical platforms have emerged to truly move the "hospital" to the mobile terminal and fully realize the non-geographical medical environment, such as Tang Ming [17] designed a wisdom rehabilitation medical service platform for the elderly and Qiu Xiaolu [18] established an Artificial Intelligence platform for autistic children, the application effects are effective and provide data and theoretical basis for the platform establishment. However, the construction of application platform standards has not yet been unified, and the construction of a complete platform system is still the focus of research.

Internet medical service model: The common symbols are "medical service system", "health consultation", "precision medicine", "outpatient service", etc. Medical service itself is a demand service, and the construction of the system has more research results in the fields of graded treatment, chronic disease management, and medical care integration. Personalized precision medicine services are designed to address the need for continuous medical and health services at the individual level, and the precision medicine strategy is part of China's 13th Five-Year
Plan, and the state plans to provide 60 billion yuan of financial support for precision medicine development by 2030 [19], strengthening the construction of precision medicine hardware and software in China and providing truly human-centered services are the keys to promoting the steady development of health care.

3.7 Evolutionary Path of Internet Healthcare Theme Hotspots

The time zone view can show the evolution trend of knowledge in Internet medical field in time dimension. 15 keywords with the highest frequency each year are used to draw the knowledge map in time zone view (Fig.6), and each circle in the cumulative time zone map represents a keyword, and the larger the circle is, larger circles indicate higher keyword frequency. China has been exploring in the field of Internet medical care, during 2011-2014, domestic Internet medical research was in the initial stage, and the keywords that appeared more frequently were mainly Internet+, mobile medical, medical equipment, etc., and the research direction was relatively macro; during 2015-2019, domestic Internet medical research studies increased rapidly and became more specific, mainly focusing on Internet hospitals, graded diagnosis and treatment, telemedicine, medical information security, medical platforms and medical health, and the application of Internet information technology in the medical field continued to deepen and the market scale gradually expanded; since 2020, the new risk development stage has seen the emergence of keywords such as online consultation and online medical care. Under the influence of the new pneumonia epidemic, people have started to have new demands for medical services, and the advantages of Internet medical care during the new pneumonia prevention period are highlighted by its convenient process and spatial freedom, and the contactless "Internet + medical health" service model has become a new trend.

![Figure 6. Keyword Cumulative Time Zone Evolution Diagram](image)

4 CONCLUSIONS

The article is based on the literature in China's Internet medical field as the research object, by using COOC12.6 and Vosviewer visualization tools to analyze important journals, major research institutions, core author groups and research hotspots in the research field, and draws the following conclusions:
First, the annual publication volume in China's Internet medical field in the past 10 years is generally on the rise, with a large number and wide range of published journals, but fewer high-quality core papers, and the research level and research depth need to be improved; the cooperation of core researchers is generally scattered, and the regional cooperation of research institutions is low; the content of high-frequency cited literature is focused on the application service model, industry development trend and the construction of the regulatory system.

Second, through keyword co-occurrence network mapping and cluster analysis, it can be seen that the hot keywords are Internet+, Internet hospital, mobile medical, medical services, etc., which are in the center of the co-occurrence network mapping and closely related to other keywords; a total of five keyword clusters are obtained, and the main research directions of their formation can be classified into the following three aspects: Internet medical information construction, Internet medical platform application and the construction of Internet medical service model, and the hot research themes have a certain stage continuity, on the one hand, attributed to the development of the information age, on the other hand, due to the impact of the new crown epidemic, all deepen the process of China's Internet medical career research in different degrees.

Third, from the evolution of subject terms, the research path in the field of Internet healthcare in China is divided into three stages: Active exploration phase (2010-2014) The research content focuses on the application of Internet technology in the medical field, exploring and applying mobile medical devices, and laying the foundation for the subsequent technological innovation of Internet medicine; Vigorous development stage (2015-2019) Research content tends to explore the construction of Internet hospital informatization and personalized medical services (such as chronic disease management, graded diagnosis and treatment, medical and health care integration, etc.). The continuous integration of health big data and artificial intelligence technology, exploring more accurate diagnosis and treatment services has become a hot topic in this period; New Risk Development Phase (2020-now) Under the influence of the new crown epidemic, the research content gradually shifts to the construction of a service system that focuses on the needs of the audience, online medical care and online consultation become hot keywords, sharing medical resources without geography and providing real-time medical service model are the frontier and direction of future Internet medical development.

In summary, the further construction of information technology for Internet hospitals and demand-driven medical services are likely to be the main directions of development in the field of Internet healthcare in the future. There are also shortcomings in the present study. The data sources are mainly Chinese literature published in journals, which has limited coverage and different parameter settings may lead to different results. In future studies, the author will continue to expand the sample database and continue to focus on the development of the Internet healthcare field.

REFERENCES

Analysis of College Students’ Consumption Structure and Irrational Consumption Behavior Based on ELES Model

Xucong Hu1,a, Weihong Wang1, Sirui Huo1, Fengzhi Chen1
huxucong0237@163.com
Faculty of Psychology, Southwest University, Chongqing, China1

Abstract: The Internet marketing industry is of rapid development currently. With the characteristics of convenience and commodity diversification, college students are an important consumer group. Mental accounting refers to the fact that people often divide into several independent accounts for different categories when making consumption decisions. Due to the poor self-control abilities, college students are prone to irrational consumption behaviors under specific e-commerce promotion festivals such as “Double 11”, in which their mental accounting structure is also changed. This study conducts a questionnaire survey on the consumption structure of college students daily and e-commerce festivals. Using ELES model analysis, it is found that the expenditure of college students’ mental accounts is mainly divided into survival, developmental and enjoying. Further, it is found that the essence of irrational consumption behavior is to squeeze survival, developmental consumption to increase enjoying consumption under the stimulation of online consumption festival. The data and results could be a reference for college students with important preventive significance.

Keywords: ELES model, College Student, Consumption, Mental Accounting.

1 INTRODUCTION

With the prospects of Internet e-commerce industry, the network consumption industry chain further improved. The network consumption platform products become various currently, the influence of “Internet celebrity economy” and e-commerce live broadcast is increasing under the new media environment, showing a blowout development trend [19]. The essence of shopping carnival is the product of the prevailing consumerism era: through the celebrity endorsements, shopping festivals and other ways to stimulate people’s desire to purchase, entice blind and irrational consumption [1]. In the post-epidemic economic downturn, online e-commerce has become the preferred choice to stimulate the economic development [18]. The shopping festival with Interactivity and entertainment complement with live broadcast, has long been transformed from new things to what public is familiar with and widely involved in.

As the young generation of consumer group with great potential, undergraduates hold a high acceptance of online shopping [17]. Due to the low threshold of online shopping and mental conformity, they are easy to listen to Internet celebrity, leading to excessive and blind consumption behaviors. They often uses “discount” as an excuse for their own impulse
consumption. The average annual expenditure of college students in China increased significantly, also with the rising of purchasing high-end and luxury products [6].

The mental accounting theory was first discovered by Thaler (1980). Influenced by mental accounting, individuals often violate some simple economic algorithms when making decisions, thus inducing many irrational consumption behaviors, which is one of the important reasons for the irrational consumption behavior of college students [3]. This is related to the non-substitutability, flexibility and other characteristics of mental accounting [5]. Flexibility refers that people tend to construct a special account to accommodate vague categories of consumption items, and when it’s tempting, they would fill it with other surplus accounts to complete the transaction [2]. Only from the perspective of consumer spending, this feature is more evident in college students’ consumption decisions (Yin, 2014). In order to further understand the structure of mental accounting expenditure during irrational consumption under specific stimuli, we pre-released and collected 164 questionnaires at the “618” consumption festival. After data analysis, we summed up two features: one is the psychological assumptions and the actual mental account does not match. That is, the difference between the expected consumption expenditure structure and the actual detailed expenditure is significant. The second is the weak execution. Those college students were vulnerable to external influences, highly likely to result in excessive consumption. During the e-commerce festival, 21% cost more than one month’s full living expense, and 47% cost more than half, but only 46% of the students believed to have irrational consumer behaviors in “618”. Most college students have a weak sense of consumption and are in a serious but unconscious dangerous state of excessive consumption. In addition to e-commerce platforms, the emergence of credit platforms with lower thresholds has led to the borrowing of college students who are not strong-willed and impulsive. This meet the economic needs of many college students, and the majority have fallen into an embarrassing dilemma [4]. Ant Huabei released “The reports of Chinese young adults daily-life consumption”, which showed every 1 in 4 youngsters would use advanced consumption on average [21]. If college students fail to repay their arrears on time, they would leave a record of dishonesty, or even fall into the trap of online loan and cannot extricate themselves, arousing more psychological problems to extreme or degenerate [10].

This study would explore the specific dimensions of college students’ mental accounts and the essential causes behind the promotion of those e-commerce festivals. For individual development, the university stage is not only a critical period to learn professional knowledge, but also an important period to learn financial management. If we could cultivate college students’ ecological consumption concept according to the explicit mental accounting, guide them to get rid of the negative influence of fetishism, liberate them from the joy of indulging in material wealth and improve their own quality development. It would help to cultivate college students’ personality and correct consumption concept. It is more conducive to the formation of green consumption patterns and the promotion of healthy consumption concepts in the whole society [14].

2 STUDY 1

Researchers have discussed the division of the implicit structure of mental accounting. For example, Kivetz (1999) divided mental accounting into regular income and windfall gain
according to consumption income. Thaler (1999) divided mental accounting into two mental accountings: common expenditure and luxury expenditure according to consumption expenditure. Some other researchers explored the Chinese mental accounting system, as Li, Ling, Fang and Xiao (2007) divided the Chinese mental accounting into three dimensions: the source, the expenditure and the storage of wealth. However, college students are a special consumer group, they have almost no independent economic access, their income source is often single but stable, mainly from their parents as living expenses. In Wu’s (2009) survey on college students’ consumption status: 96% of their consumption sources are dominated by the living expenses provided by their parents. Similar results were found in the survey conducted by Deng et al. (2005) that 95% of college students depend on their parents (families) for daily consumption. However, their consumption expenditure types are diverse, and an effective way to classify consumption expenditure structure is based on ELES model.

Extend Linear Expenditure System (ELES) is a demand function system introduced by C.Liuch (1973) based on the linear expenditure system model. The model assumes that people's demand for goods is divided into basic demand and marginal demand, and has nothing to do with income. On the basis of meeting basic needs, the remaining marginal income will be distributed to other aspects of life. It has the advantages of complete theoretical basis, small data dependence and easy parameter estimation. It overcomes the limitations of Engel coefficient and can digitally reflect the consumption structure of college students.

Assuming that people’s consumption expenditure is specifically divided into i types, the consumption expenditure of various commodities can be expressed by the model as follows:

$$V_i = P_i X_i + \beta_i (Y - \sum P_i X_i), \quad i = 1, 2, 3, ..., n$$ (1)

Among them, i is the type of the good or service. $V_i$ is the consumption expenditure of i good, $P_i$ and $X_i$ are the price and basic demand of i good, $P_i X_i$ is the basic demand expenditure of i good, $\beta_i$ is the marginal consumption tendency, and $Y$ is the disposable income.

To deform the upper equation:

$$V_i = \beta_i Y + (P_i X_i - \beta_i \sum P_i X_i), \quad i = 1, 2, 3, ..., n$$ (2)

Let

$$P_i X_i - \beta_i \sum P_i X_i = \beta_i$$ (3)

$$V_i = \alpha_i + \beta_i Y \quad (i = 1, 2, ..., n)$$ (4)

Let

$$\sum \alpha_i = \sum P_i X_i - \sum \beta_i \sum P_i X_i = \sum P_i X_i (1 - \sum \beta_i)$$ (5)

So we can get $\sum P_i X_i = \sum \alpha_i/(1 - \sum \beta_i)$. Thus, the basic demand for i good is:
\[ P_i X_i = \alpha_i + \beta_i \left[ \frac{\sum \alpha_i}{1 - \sum \beta_i} \right] \]  

Model (4) is a simple linear regression model. Using the cross-section material, the estimated value of \( \alpha \) and \( \beta \) can be obtained by the least-square method.

Based on the ELES model, Ju (2014) conducted an empirical survey of several universities in Shanghai, and divided the consumption expenditure structure into basic living needs, developmental needs and enjoyment needs. On the basis of this, we speculate that the expenditure dimension of mental accounting is divided into three parts based on different types of consumption purposes: survival consumption (Mainly includes the most basic food, accommodation and expenses for the basic survival guarantee), developmental consumption (Mainly includes personal hobbies and individual developmental needs), enjoying consumption (Mainly includes games, love and other entertainment-oriented consumption purposes).

2.1 Methods and Materials

2.1.1 Participants

The participants were college students from Southwest University, in Chongqing, China. 416 data were distributed and recovered, in which there was 403 valid data, with an effective rate of 96.88%. There were 189 males (46.90%) and 214 females (53.10%), and mean age is 20.72 years old.

2.1.2 Materials

The “College Students Daily Consumption Expenditure Questionnaire” used in this study is mainly compiled, based on Ju’s (2014) questionnaire on the consumption status of college students. The questionnaire mainly focuses on the consumption income and sources of college students, and summarizes their common consumption expenditure categories, including: basic food, basic learning, transportation, accommodation, self-development, social contact, love, entertainment, image consumption, enjoying dietary, and further investigates the proportion of their daily consumption.

2.2 Results

2.2.1 Description Statistics of College Students Daily Consumption Structure

As shown in Table 1, we found that 98.01% of college students’ income comes from the living expenses provided by their parents, while only 11.41% had additional personal fixed income and 5.70% had other sources of income. It could be seen that for college students, the main source of consumption income remained the living expenses from their parents.

The average monthly income of college students were divided into 4 levels from 0 to 3000 yuan. The analysis results showed that the average monthly living expenses were about 2051.42 yuan, and the college students whose monthly living expenses were in the range of 1000-2000 occupied the main part, reaching 68.98%, followed by 20.84% whose monthly living expenses were in the range of 2000 - 3000. While those in the range of 0-1000 accounted for 6.69% and those above 3000 accounted for 3.72%.
Table 1: Consumption income and source of college students.

<table>
<thead>
<tr>
<th>Content</th>
<th>Group</th>
<th>Number</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source</td>
<td>Living expense from parents (family)</td>
<td>395</td>
<td>98.01%</td>
</tr>
<tr>
<td>Part-time income</td>
<td></td>
<td>46</td>
<td>11.41%</td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td>23</td>
<td>5.70%</td>
</tr>
<tr>
<td>Monthly income (yuan)</td>
<td>0 - 1000</td>
<td>27</td>
<td>6.69%</td>
</tr>
<tr>
<td></td>
<td>1000 - 2000</td>
<td>278</td>
<td>68.98%</td>
</tr>
<tr>
<td></td>
<td>2000 - 3000</td>
<td>84</td>
<td>20.84%</td>
</tr>
<tr>
<td></td>
<td>&gt; 3000</td>
<td>15</td>
<td>3.72%</td>
</tr>
</tbody>
</table>

The average monthly expenditure was about 1622.5 yuan, but the analysis showed that 44.34% of undergraduates’ monthly consumption expenditure outweighed their monthly living expenses. The specific proportion and structure of different types of consumption are further counted (See in Figure 1). The proportion of basic food, basic learning, transportation, accommodation, social contact, self-development, love, entertainment, clothing, makeup and enjoying dietary in monthly expenditure is: 35.36%, 5.03%, 4.37%, 3.85%, 6.52%, 6.12%, 3.44%, 9.38%, 11.50%, 3.54%, 9.89%. According to the different consumption purposes, we could find that survival consumption accounted for the highest proportion, reaching 48.61%, while developmental consumption and enjoying consumption accounted for 12.63% and 34.31% respectively.

Figure 1: Consumption expenditure categories of college students.

2.2.2 ELES Analysis of College students Daily Consumption Structure

SPSS 21.0 was used to analyze the data obtained from the “College Students Daily Consumption Expenditure Questionnaire”. Taking the cross-sectional data of college students’ average monthly income and consumption expenditure as the data source, the ELES model was used to carry out regression analysis on their consumption structure, in which income was the independent variable and expenditure was the dependent variable. The statistical results are shown in Table 2. From the determination coefficient, t test and F test value, we could find the fitting effect of ELES model is great.
According to the results shown in Table 2, from the marginal consumption tendency \((\beta)\) of various types of consumption, the marginal consumption tendency of college students from high to low is entertainment, basic food, clothing, social contact, enjoying dietary, basic learning, makeup, transportation, self-development, love, accommodation. Among them, the marginal consumption tendency of entertainment, basic food and clothing is greater than 0.1, which shows that with the increase of living expenses, college students would spend more on entertainment and image consumption, and have higher requirements on the quality of daily food. In contrast, the marginal consumption tendency of transportation, accommodation and self-development is low, indicating that the improvement of campus infrastructure lifts the supply constraints and the willingness of college students to develop themselves is small.

Table 2: Parameter Estimation and Statistical Test Results of ELES Model.

<table>
<thead>
<tr>
<th>Consumption Categories</th>
<th>(\alpha)</th>
<th>(\beta)</th>
<th>(R^2)</th>
<th>(t)</th>
<th>(F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic food</td>
<td>469.317</td>
<td>.138</td>
<td>.073</td>
<td>2.866</td>
<td>8.214</td>
</tr>
<tr>
<td>Basic learning</td>
<td>-33.520</td>
<td>.079</td>
<td>.079</td>
<td>2.978</td>
<td>8.868</td>
</tr>
<tr>
<td>Transportation</td>
<td>-16.124</td>
<td>.060</td>
<td>.190</td>
<td>4.946</td>
<td>24.464</td>
</tr>
<tr>
<td>Accommodation</td>
<td>81.266</td>
<td>-.001</td>
<td>.010</td>
<td>-.048</td>
<td>.002</td>
</tr>
<tr>
<td>Social contact</td>
<td>.550</td>
<td>.081</td>
<td>.161</td>
<td>4.462</td>
<td>19.907</td>
</tr>
<tr>
<td>Self-development</td>
<td>37.750</td>
<td>.053</td>
<td>.061</td>
<td>2.591</td>
<td>6.713</td>
</tr>
<tr>
<td>Love</td>
<td>10.046</td>
<td>.038</td>
<td>.034</td>
<td>1.926</td>
<td>3.708</td>
</tr>
<tr>
<td>Entertainment</td>
<td>-108.189</td>
<td>.186</td>
<td>.134</td>
<td>4.017</td>
<td>16.135</td>
</tr>
<tr>
<td>Clothing</td>
<td>27.099</td>
<td>.125</td>
<td>.106</td>
<td>3.511</td>
<td>12.324</td>
</tr>
<tr>
<td>Makeup</td>
<td>-14.470</td>
<td>.067</td>
<td>.144</td>
<td>4.186</td>
<td>17.526</td>
</tr>
<tr>
<td>Enjoying dietary</td>
<td>60.279</td>
<td>.086</td>
<td>.111</td>
<td>3.596</td>
<td>12.932</td>
</tr>
</tbody>
</table>

Subsequently, the equation for calculating the basic needs based on the ELES model is:

\[
P_iX_i = \alpha_i + \beta_i \left[ \frac{\sum \alpha_i}{1 - \sum \beta_i} \right]
\]  

(7)

For example, the basic demand for basic food is \(P_1X_1 = 540.12\), and the equation for income elasticity of demand is:

\[
n_i = \beta \cdot iY / V
\]

(8)

Further statistical analysis is made on the estimation of college students’ basic consumption demand and income elasticity of demand. The results are shown in Table 3. The data shows that the valuation of most consumption categories is positive, and only the valuation of entertainment is negative, indicating that after the basic consumption demand of college students is met, the remaining income is more inclined to be used for entertainment. At the same time, according to the calculation results of the demand income elasticity equation, the income demand elasticity of entertainment and makeup is higher, which represents with the
increase of the income level of college students, they are more inclined to invest in the enjoying consumption related to entertainment and image.

<table>
<thead>
<tr>
<th>Consumption Categories</th>
<th>Basic consumption demand forecast</th>
<th>Income elasticity of demand</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic food</td>
<td>540.12</td>
<td>0.33</td>
</tr>
<tr>
<td>Basic learning</td>
<td>7.01</td>
<td>0.74</td>
</tr>
<tr>
<td>Transportation</td>
<td>14.66</td>
<td>0.89</td>
</tr>
<tr>
<td>Accommodation</td>
<td>80.75</td>
<td>-0.02</td>
</tr>
<tr>
<td>Social contact</td>
<td>42.11</td>
<td>1.00</td>
</tr>
<tr>
<td>Self-development</td>
<td>64.94</td>
<td>0.70</td>
</tr>
<tr>
<td>Love</td>
<td>29.54</td>
<td>0.86</td>
</tr>
<tr>
<td>Entertainment</td>
<td>-12.75</td>
<td>1.54</td>
</tr>
<tr>
<td>Clothing</td>
<td>91.24</td>
<td>0.88</td>
</tr>
<tr>
<td>Makeup</td>
<td>19.91</td>
<td>1.15</td>
</tr>
<tr>
<td>Enjoying dietary</td>
<td>104.40</td>
<td>0.70</td>
</tr>
</tbody>
</table>

2.3 Discussion

We analyzed the income sources and expenditure categories of college students’ consumption, and found that consistent with previous studies, the majority of the college students’ income sources are still living expenses, and mainly concentrated in the range of 1000-3000, but the phenomenon of excessive consumption still exists. Based on the classification of mental accounts, it could be found that most of the consumption in college students’ life is still survival consumption, such as basic food and accommodation, followed by enjoying consumption, while less college students invest in consumption for developmental purposes. In the subsequent ELES model analysis, it is also found that with the increase of living expenses, college students mainly invest more in enjoyable expenditure, such as entertainment, image, etc., indicating that for college students, the most attractive type of consumption is the mental account of enjoyment, which is also the cause of irrational consumption. It would be analyzed later.

3 STUDY 2

Consumption is a social interactive behavior, which often starts with the individual’s motivation to meet the need of consciousness or emotional experience. Therefore consumer behavior is a manifestation of the internal process to satisfy self-demand, under the initiation of external stimuli (inducement) (Wang, 2003).

The “online consumption festival” (e.g. “618”, “double 11”) initiated by the online shopping platform is often accompanied by activities in the form of discounts, promotions and so on, therefore consumers are favored by price advantages. This kind of external stimuli stimulates
consumers’ consumption motivation and makes it easier for them to make consumption decisions. In the eyes of consumers, “online consumption festival” means lower prices, and also the consumption time with most cost-efficiency, the inherent concept of “missing this time is a loss” is generated. In the long run, it forms a conditional reflection. Without comparing the price changes or paying attention to self-actual consumption needs, they would directly regard those festivals as the best time for consumption, resulting in the negative schema of “consume for consumption”. College students are still in the prefrontal development stage, so their emotional control and rational decision-making functions still immature. Once there is an external stimulation as “online consumption festival”, it is easier for college students to make decisions on enjoying consumption (e.g. entertainment, high-end clothing, makeup). However, the monthly consumption income of college students is fixed and controlled by the parents. When due to the fixity, in order to ensure the balance of income and expenditure, college students are often unconsciously affected by the independence of mental accounts, and unconsciously reduce their expenditure on survival consumption and developmental consumption. According to survey by Jiang (2022), there are no less than 100 various sizes of promotional activities in Taobao platform within one year. The increase of enjoying consumption after the stimulation functions as a reward for consumers. They cost only a few more every time, but the spiritual satisfaction feedback continues to strengthen the consumption imbalance of college students. Over time, when the enjoying consumption gradually squeeze survival consumption so that the most basic survival needs could not meet (e.g. food, clothing, etc.). This would be detrimental to health, and even induce borrowing and other advanced consumption behaviors (See the specific psychological mechanism in Figure 2).

According to the analysis of the psychological mechanism, it is predicted that college students would significantly increase the proportion of enjoying consumption, reduce the proportion of survival consumption and developmental consumption, and even exceed the monthly disposable income like overdrawing next month’s part, during the “online consumption festival” to keep the balance.

![Figure 2: Psychological mechanism of irrational consumption.](image)

3.1 Methods and Materials

3.1.1 Participants

The participants were also college students from Southwest University, in Chongqing, China. 406 data were distributed and recovered, in which there was 399 valid data, with an effective rate of 98.28%. There were 194 males (48.62%) and 205 females (51.38%), and mean age is 20.43 years old.
3.1.2 Materials

On the basis of the “College Students Daily Consumption Expenditure Questionnaire”, we changed the guidance, and compiled the “College Students Consumption Expenditure Questionnaire on Online Consumption Festival”. It was released in the end of November after the “Double 11 Consumption Festival”, focusing on not only the category and specific amount of consumption expenditure in the past week’s consumption festival, but also counting normal expenses this month.

3.2 Results

3.2.1 College Students' Consumption Structure on “Online Consumption Festival”

According to the results, 95.74% of the participants had consumption behavior in the “Double 11 Consumption Festival”, and found that the average consumption in the festival reached 1358.84 yuan, which accounted for about 64.39% of the monthly living expenses. However, more than 16.63% of college students would consume more than the monthly living expenses in the festival, and the highest could even reach 7.47 times of the monthly average disposable income.

The results are consistent with the hypothesis. As shown in Figure 3, the consumption of college students in the online consumption festival is mainly enjoying consumption, including enjoying dietary, clothing, bag, makeup, daily necessities, digital products. Clothing and enjoying dietary accounted for the largest proportion, reaching 28.77% and 23.63%, followed by makeup and digital products accounted for 15.43% and 18.76%, daily necessities and bags accounted for 8.87%, 2.43%.

![Figure 3: Detailed expenditure categories on “Double 11 Festival” (From high to low).](image)

3.2.2 Comparative Analysis of Daily and “Online Consumption Festival” Consumption Structure

In the results, we analyzed the proportion of college students’ daily expenditure on survival, developmental and enjoying consumption, and calculated the consumption structure after the “Double 11 Consumption Festival” to obtain the specific proportion of three types of consumption. The comparison results are shown in Figure 4. It could be found that under the influence of the festival, the weight of survival consumption and developmental consumption decreased, while the weight of enjoying consumption increased significantly. In order to
further judge whether the difference was significant, we would use variance analysis to observe weight changes.

![Figure 4: Consumption structure in daily and online consumption festival.](image.png)

In the two-factor analysis of variance, with time (daily vs online consumption festival) and mental accounting type (survival vs developmental vs enjoying consumption) as independent variables and consumption weight as dependent variable, it was found that the main effect of mental accounting type was significant \((F = 105.224, p < 0.01)\), indicating that there were significant differences among the three different consumption types, but the main effect of time and the interaction between time and consumption type were not significant \((p > 0.05)\). The further pairwise comparison results showed that the difference of survival consumption between daily and the online consumption festival was significant \((t = 12.179, p < 0.01)\), the difference of enjoying consumption was also significant \((t = 26.151, p < 0.01)\), while the difference of developmental consumption was not significant \((p > 0.05)\). It could be found that the survival consumption of college students after the festival significantly reduced, the enjoying consumption significantly increased, and the development consumption nearly unchanged. This may due to the large individual difference in the direction of personal development, while the survival and enjoying consumption changes basically met the hypothesis.

### 4 GENERAL DISCUSSION

This study mainly used ELES model to analyze the consumption structure of college students’ mental accounts and finds three types: survival, developmental and enjoying. On this basis, through the comparative analysis of the changes in the structure of mental accounts by college students before and after the “Double 11” consumption festival, the reasons for their irrational consumption behavior of college students are explored. It is found that its essence is to compress the survival and developmental consumption to carry out enjoying consumption, which reflects the flexibility of mental accounts and would bring negative consequences to college students for a long time, such as affecting physically healthy development and even online lending. The theoretical significance of this study is to explore the consumption structure and mental accounting dimension of college students, and use ELES model analysis
to obtain data as supporting evidence. From this perspective, the causes of irrational consumption behavior were studied. In the meanwhile, it also has practical significance. Through the analysis, it provides reference for college students to understand their consumption proportion and tendency, so as to effectively control self-consumption and prevent potential irrational consumption behavior.

However, there are still some limitations in this study, which is to be further expanded in future. First of all, the sample size needs to be further expanded. All the participants invited are from Southwest University, which may only represent the consumption structure of college students in specific region (Southwest). In order to better explore the general consumption structure of Chinese college students, we should combine the consumer price index of each region, and use statistical methods like factor analysis in future (Lin, 2012). Secondly, since the self-report form of the questionnaire is mainly used, it is impossible that some college students could remember every consumer spending monthly without missing. Therefore, a better alternative approach is to monitor students’ monthly expenditure under the stimuli of online consumption festival, with the premise of informed consent. In this way could the accurate data be obtained. In addition, another future research direction is to put these data obtained through the ELES model into real life into intervention, which could flexibly calculate the reference value of the month according to individual income, so as to play a good preventive role in irrational consumption behavior.

5 CONCLUSIONS

To review, with the analysis of ELES model, this paper explores and verifies that there are three dimensions of college students’ mental accounting: survival, developmental and enjoying consumption. They are independent of each other, but under the stimulation of online consumption festival, college students often could not help the pleasure brought by enjoying consumption, resulting in irrational consumption behaviors, which in turn squeezes survival and developmental consumption, threatening to their long-term healthy development.

Acknowledgements: This research was supported by grants from Weihong Wang, and the SWU Research Funds (S202210635294) to Xucong Hu.

REFERENCES

contiguity in consumer-borrowing decisions. Organizational Behavior and Human Decision Processes, 58(1), 136-152.
Analysis of influencing factors of energy data market construction and development based on entropy weight DEMATEL method

ChunLei Zhou¹, JunNi Li¹, TianGuang Yang², JinWei Song¹, Yu Zhang², SiYuan Gao³*, YaFeng Wen³

*Correspondence: 15653700817@163.com

Big Data Center, State Grid Corporation of China, Beijing 100052, China¹
Beijing SGITG-Accenture Information Technology, Beijing 100052, China²
School of Electrical and Electronic Engineering of the North China Electric Power University, Beijing 102206³

Abstract. The construction of data market is of great strategic significance to the development of energy industry. Aiming at the key influencing factors of the construction and development of energy data market, we analyze the influence of policies and laws, market system, value creation and big data technology on energy data market, and construct an evaluation index system of influencing factors of energy data market. Then, entropy weight -DEMATEL method is used to comprehensively evaluate each index. Finally, the evaluation results of the index system are evaluated by experts, and it is concluded that the most basic factors affecting the construction and development of energy data market are business compliance, social benefits of products, environmental benefits of products, data product development, data sharing technology and data privacy protection.

Keywords. energy data market; Data transaction; Energy data; Data assets; Entropy weight -DEMATEL method;

1 Introduction

At present, the data trading mechanism is the research hotspot. Literature [1], aiming at data trading, proposes a data trading mechanism oriented to federated learning, which provides an effective means to get rid of the dilemma of "data island". Literature [2] establishes a data market based on auction mechanism to promote the interaction between renewable energy production and forecast data. Literature [3] proposes a pricing mechanism that comprehensively considers data contribution, integrity and query quality. Literature [4] proposes a data resource pricing method based on the multi-algorithm fusion model for Stacking in view of the information asymmetry between buyers and sellers in data transactions. Literature [5] divides data ownership according to data type and subject classification. Bao Xiaoli [6] also considered the process of data value generation and the legal attributes of data, and proposed a second-order sequential data ownership rule. Literature [7] designed a trusted data trading ecosystem based on blockchain. In order to explore the key influencing factors of
the construction and development of the energy data market, establishes the evaluation index system of the influencing factors of the energy data market. Then, the entropy weight -DEMATEL method is used to evaluate the expert score results, and the comprehensive situation of the influencing factors of the energy data market is obtained. According to the research results, six indexes, namely business compliance, product social benefit, product environmental benefit, data product development, data sharing technology, and data privacy protection, are the main influencing factors for the construction and development of energy data market, and suggestions are put forward accordingly.

2 Construction of index system of energy data market influencing factors

2.1 Analysis of influencing factors of energy data market

2.1.1 Valence dimension selection

At the early stage of the construction of the energy data market, the policy environment plays a guiding role in the market activities. The legal environment regulates the market behavior, which is the basic guarantee for the operation and development of the energy data market. Data trade produces direct economic benefits, and data circulation can bring indirect economic benefits. The construction of energy data market is not only to promote the flow of data itself, but also to promote the flow of energy data value. To sum up, this topic from the policy and legal environment, value creation, market management system, data flow technology, the main influencing factors of energy data market are analyzed from four aspects.

2.1.2 Policy and Legal Environment

In recent years, the transformation of digital strategy has become an important strategy of all countries in the world, and many countries and regions have introduced laws and regulations to promote data transaction. Our country, the "14th Five-Year Plan for Digital Economy Development", the "14th Five-Year Plan for Digital Economy Development". Relevant supporting laws and regulations are also being gradually introduced to further promote the formal development of the data factor market. On the one hand, data transactions need to be regulated and restricted by laws. Data ownership, data transactions and data privacy protection all need to be based on laws. Legal transaction compliance is the most basic guarantee of data market. At present, there is no normative document with legal effect aiming at the big data trading activities. On the other hand, policy support is an important power source for the construction of energy data market.

2.1.3 Market management system

Through the energy data market management system and the formulation of the code of conduct of participants, the standardization of the whole life cycle of data trading can be realized, with the main rights confirmation, pricing, trading and supervision systems. Data confirmation is the basic condition of data transaction. The confirmation of data ownership is very complicated. In addition, the establishment of property rights of data commodities should consider not only legal attributes, but also market factors. Data pricing mechanism can
standardize the market behavior of data trading fundamentally, and data trading mechanism is the basis of data trading. Establishing and improving the trading mechanism of data commodities is a necessary condition for constructing a complete energy data market. Market transaction regulation is also an essential link in energy data trading. The development of efficient and reliable data asset management system can promote the transformation of energy data resources into data assets. In the aspect of data supply management, data collection and transmission mechanism is a necessary condition for the stability of data market supply. To sum up, from the dimension of data market management system, we select data asset management, data right confirmation system, data circulation system, data pricing system, data trading system, data trading supervision, data operation strategy, data reliable collection, and data efficient access as influencing factors.

2.1.4 Value creation

As the product of the combination of energy and big data, energy data will further strengthen the connection between energy and information, improve energy efficiency on the one hand, and increase the value of information transmission on the other hand. It is not only of great strategic significance to the development of the energy field, but also of great value to the society, economy and environment. Energy data trading can strengthen the cooperation between government and enterprises, promote the collection and sharing of data in the field of energy and public services, and form a powerful social governance force. The establishment of energy data market creates huge value for the government, enterprises and other data subjects, which will promote the construction of data market and form a virtuous cycle.

2.1.5 Big Data Technology

Data trading needs to rely on big data technology, which mainly includes data management and processing technology, data sharing technology, data fusion technology and data product development technology. Data access, storage and access control require the support of data management and processing technology. Data sharing technology mainly aims at the security and sharing efficiency in the process of data transaction. Safe and reliable data sharing environment and perfect privacy protection mechanism are the basic guarantee for data transaction. In the energy data market, there are a wide range of data sources and big data differences, leading to big differences in model standards, design specifications, coverage areas and other aspects, resulting in difficult data aggregation. Therefore, it is necessary to carry out data fusion, so as to lay the foundation for data sharing and trading. Data product development technology is the basic support of data assets.

2.2 Analysis of influencing factors of energy data market

Based on the above research content, an energy data market evaluation index system, including 24 secondary indexes, is constructed from four dimensions of policy and law, market system, value creation and big data technology, as shown in Table 1.
Table 1 index system of influencing factors of energy data market

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Sub dimension</th>
<th>Influence factor</th>
<th>The serial number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Policy and law</td>
<td>Laws and regulations</td>
<td>Business compliance</td>
<td>A1</td>
</tr>
<tr>
<td></td>
<td>Policy</td>
<td>Level of policy support</td>
<td>A2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Policy orientation</td>
<td>A3</td>
</tr>
<tr>
<td>The market system</td>
<td>Data management system</td>
<td>Data Asset Management</td>
<td>B1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Data confirmation system</td>
<td>B2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Data flow system</td>
<td>B3</td>
</tr>
<tr>
<td>Data transaction system</td>
<td></td>
<td>Data pricing system</td>
<td>B4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Data transaction system</td>
<td>B5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Data transaction supervision</td>
<td>B6</td>
</tr>
<tr>
<td>Data supply incentive</td>
<td></td>
<td>Data Operation Strategy</td>
<td>B7</td>
</tr>
<tr>
<td>Value creation</td>
<td>Social value</td>
<td>Social benefits</td>
<td>C1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Environmental benefits</td>
<td>C2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Industry benefits</td>
<td>C3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Technical efficiency</td>
<td>C4</td>
</tr>
<tr>
<td>The economic value</td>
<td></td>
<td>Data transaction cost</td>
<td>C5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Direct economic benefits</td>
<td>C6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Indirect economic benefit</td>
<td>C7</td>
</tr>
<tr>
<td>Big Data Technology</td>
<td>Data processing technology</td>
<td>Data management Technology</td>
<td>D1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Data processing capacity</td>
<td>D2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Data Product Development</td>
<td>D3</td>
</tr>
<tr>
<td>Data flow technology</td>
<td></td>
<td>Data sharing technology</td>
<td>D4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Data privacy protection</td>
<td>D5</td>
</tr>
</tbody>
</table>

3 Construction of index system of energy data market influencing factors

In this paper, entropy weight -DEMATEL method is used to comprehensively evaluate the scoring results of the index system. Considering the complex relationship between different factors, DEMATEL method was introduced to quantify the interaction between different factors based on the entropy weight method. The evaluation method based on entropy weight -DEMATEL can not only overcome the subjectivity of expert rating to a certain extent, but also prevent the internal relationship of influencing factors from being ignored, so as to reduce
the evaluation error.

3.1 Entropy weight method to determine the index weight

Firstly, the original evaluation matrix is normalized. On this basis, the entropy weight method is used to determine the weight of each index. The entropy $e_j$ of the JTH index calculated by the entropy weight method is:

$$e_j = -\frac{1}{\ln m} \sum_{i=1}^{m} p_{ij} \ln p_{ij}$$

(1)

$$p_{ij} = z_{ij} / \sum_{j=1}^{n} z_{ij}$$

(2)

$p_{ij}$ is the proportion of the ith evaluation scheme of the JTH evaluation index. The dispersion degree $d_j$ of the evaluation data of the JTH index is expressed as $d_j = 1 - e_j$. Therefore, the weight $w_j$ of the JTH index is expressed by the entropy weight method as:

$$w_j = d_j / \sum_{j=1}^{n} d_j = (1 - e_j) / \sum_{j=1}^{n} (1 - e_j)$$

(3)

3.2 DEMATEL method was used to calculate impact

DEMATEL method is used to analyze the logical relation and direct influence matrix of each influencing factor. Based on graph theory and matrix operation, the effects of each element on other elements are found out, and the relationship between each element and the position of each element in the index system are found out. To analyze impact using the DEMATEL method, perform the following steps.

1) The direct influence matrix $M$ is obtained by quantifying the relationship between various influencing factors.

$$M = \begin{bmatrix}
0 & a_{12} & \cdots & a_{1n} \\
a_{21} & 0 & \cdots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{n1} & a_{n2} & \cdots & 0
\end{bmatrix}$$

(4)

2) By normalizing the original relation matrix, the norm which directly affects the matrix $N$ is obtained.
$$M_i = \max \left\{ \sum_{j=1}^{n} M_{ij} \right\}$$  \hspace{1cm} (5)

$$b_{ij} = \frac{M_{ij}}{M_i}$$  \hspace{1cm} (6)

$$N = \left[ b_{ij} \right]$$  \hspace{1cm} (7)

3) The comprehensive influence matrix $T$ is obtained from the normalized direct influence matrix.

$$T = N \ast (I - N)^{-1}$$  \hspace{1cm} (8)

4) According to the values $D = (D_1, D_2, \cdots, D_n)$, $C = (C_1, C_2, \cdots, C_n)$, in the comprehensive impact matrix $t$, the impact degree $D_i$, impact degree $C_i$, centrality degree $M_i$, and cause degree $R_i$ of each factor were calculated. Impact reflects the impact of an indicator on other indicators. Impact refers to the impact of other indicators on the indicator. Centrality refers to the role of the index in the whole system. The effect of cause degree index on other factors. If the cause degree is greater than 0, other indicators will be affected by this indicator. If the cause degree is less than 0, the indicator will be affected by other indicators.

$$\begin{align*}
D_i &= \sum_{j=1}^{n} t_{ij} \\
C_i &= \sum_{j=1}^{n} t_{ji} \\
M_i &= D_i + C_i \\
R_i &= D_i - C_i
\end{align*}$$  \hspace{1cm} (9)

5) For each index, the weight calculated by the entropy weight method is multiplied by the centrality, and then normalized to obtain the comprehensive obstacle degree of each factor.

4 Analysis of influencing factors of energy data market

4.1 Analysis of evaluation results

Establish an expert review team, and obtain the index value of each business through the expert score. Among them, the quantitative method of qualitative evaluation index is shown in Table 2.
Table 2 quantification of energy data market construction indicators

<table>
<thead>
<tr>
<th>The index type</th>
<th>The worst/least</th>
<th>Poor/less</th>
<th>Good/high</th>
<th>The best/highest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benefit index</td>
<td>50</td>
<td>65</td>
<td>80</td>
<td>95</td>
</tr>
<tr>
<td>Index of the cost</td>
<td>95</td>
<td>80</td>
<td>65</td>
<td>50</td>
</tr>
</tbody>
</table>

The original matrix is normalized, and then the weight of each index is calculated by entropy weight method. The weights of each index are shown in Table 3.

Table 3 index weight of energy data market construction

<table>
<thead>
<tr>
<th>indicators</th>
<th>The weight</th>
<th>ranking</th>
<th>indicators</th>
<th>The weight</th>
<th>ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>0.076</td>
<td>1</td>
<td>D1</td>
<td>0.05</td>
<td>13</td>
</tr>
<tr>
<td>C2</td>
<td>0.068</td>
<td>2</td>
<td>D3</td>
<td>0.05</td>
<td>14</td>
</tr>
<tr>
<td>C1</td>
<td>0.063</td>
<td>3</td>
<td>B9</td>
<td>0.048</td>
<td>15</td>
</tr>
<tr>
<td>A2</td>
<td>0.062</td>
<td>4</td>
<td>D2</td>
<td>0.047</td>
<td>16</td>
</tr>
<tr>
<td>D5</td>
<td>0.061</td>
<td>5</td>
<td>C3</td>
<td>0.038</td>
<td>17</td>
</tr>
<tr>
<td>D4</td>
<td>0.059</td>
<td>6</td>
<td>B6</td>
<td>0.035</td>
<td>18</td>
</tr>
<tr>
<td>B1</td>
<td>0.053</td>
<td>7</td>
<td>C4</td>
<td>0.035</td>
<td>19</td>
</tr>
<tr>
<td>B3</td>
<td>0.052</td>
<td>8</td>
<td>B5</td>
<td>0.027</td>
<td>20</td>
</tr>
<tr>
<td>B2</td>
<td>0.051</td>
<td>9</td>
<td>B4</td>
<td>0.011</td>
<td>21</td>
</tr>
<tr>
<td>B8</td>
<td>0.051</td>
<td>10</td>
<td>B7</td>
<td>0.011</td>
<td>22</td>
</tr>
<tr>
<td>A3</td>
<td>0.05</td>
<td>11</td>
<td>C6</td>
<td>0.011</td>
<td>23</td>
</tr>
<tr>
<td>C5</td>
<td>0.05</td>
<td>12</td>
<td>C7</td>
<td>0.011</td>
<td>24</td>
</tr>
</tbody>
</table>

Fig. 1 scatter diagram of centre degree-cause degree

As can be seen from Figure 1, data circulation system, data trading system, product social benefits, product environmental benefits, data product development, data sharing technology, and data privacy protection have a great impact on the system. The index of big data
technology dimension has a high degree of cause, which has a great influence on other factors. The reason degree of benefit index is low, and it is greatly affected by other factors.

Finally, the results of entropy weight method and DEMATEL method are calculated comprehensively, and the results are sorted. The results are shown in Table 4.

<table>
<thead>
<tr>
<th>Influence factor</th>
<th>Comprehensive influence degree</th>
<th>ranking</th>
<th>Influence factor</th>
<th>Comprehensive influence degree</th>
<th>ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Business compliance</td>
<td>0.5029</td>
<td>1</td>
<td>Reliable data collection</td>
<td>0.2039</td>
<td>13</td>
</tr>
<tr>
<td>Data circulation system</td>
<td>0.3429</td>
<td>2</td>
<td>Data management technology</td>
<td>0.1896</td>
<td>14</td>
</tr>
<tr>
<td>Environmental benefit</td>
<td>0.3376</td>
<td>3</td>
<td>Data trading system</td>
<td>0.187</td>
<td>15</td>
</tr>
<tr>
<td>social results</td>
<td>0.3152</td>
<td>4</td>
<td>Data transaction cost</td>
<td>0.1801</td>
<td>16</td>
</tr>
<tr>
<td>Data product development</td>
<td>0.3087</td>
<td>5</td>
<td>Industry benefits</td>
<td>0.1585</td>
<td>17</td>
</tr>
<tr>
<td>Data Asset Management</td>
<td>0.3067</td>
<td>6</td>
<td>Technical benefits</td>
<td>0.1536</td>
<td>18</td>
</tr>
<tr>
<td>Data privacy protection</td>
<td>0.2888</td>
<td>7</td>
<td>Efficient data access</td>
<td>0.1474</td>
<td>19</td>
</tr>
<tr>
<td>Data sharing technology</td>
<td>0.2833</td>
<td>8</td>
<td>Data transaction supervision</td>
<td>0.0835</td>
<td>20</td>
</tr>
<tr>
<td>Data confirmation system</td>
<td>0.2441</td>
<td>9</td>
<td>Data pricing system</td>
<td>0.0384</td>
<td>21</td>
</tr>
<tr>
<td>Policy orientation</td>
<td>0.2372</td>
<td>10</td>
<td>Direct economic benefits</td>
<td>0.034</td>
<td>22</td>
</tr>
<tr>
<td>Data processing capacity</td>
<td>0.2223</td>
<td>11</td>
<td>Data operation strategy</td>
<td>0.026</td>
<td>23</td>
</tr>
<tr>
<td>Policy support</td>
<td>0.2076</td>
<td>12</td>
<td>Indirect economic benefits</td>
<td>0.0245</td>
<td>24</td>
</tr>
</tbody>
</table>

It can be seen from Table 4 that the key factors affecting the construction and development of the energy data market include business compliance, data circulation system, product environmental benefits, product social benefits, data product development, and data asset management, with a comprehensive impact of more than 0.3, far higher than other factors. Business compliance, data circulation system, data privacy protection, data sharing technology,
data confirmation system and other indicators related to the construction of data market system have a greater impact.

4.2 Suggestions and prospects

Based on the above evaluation and analysis, the following suggestions are made for the construction and development of the energy data market.

1) Establish and improve the data market management system

As a platform for the collection and circulation of energy data, the main purpose of the establishment of energy data trading market is to promote the circulation and sharing of data through safe and convenient trading, which requires a perfect system as the basis for the operation and development of data market. With the goal of building a unified large market for energy data that is unified and open, compliant in the circulation of data resources and the trading and circulation mechanism, the market circulation, pricing and trading mechanisms for energy data should be established, and gradually integrated into the overall market mechanism for energy data elements.

2) To explore the social and environmental value of energy data

Adhere to the demand-oriented approach, focus on tapping the national ministries, the energy industry and social enterprises' demand for energy data application, give full play to the value of data elements. By providing a series of data analysis products and services in the field of energy, continuously obtains government support and guidance, establishes a mechanism of government-enterprise joint and multi-party participation, forms a joint construction force, and collaboratively promotes the construction and operation.

3) Big data technology to support the flow of data

Big data technology is the basic support for the transformation of energy data resources into assets. On the one hand, data sharing technology should be developed to improve the efficiency and security of data trading and sharing process. On the other hand, it is necessary to introduce cutting-edge big data technology to fully tap the value of energy data.

5 Conclusion

This paper constructs an energy data market evaluation index system including 24 indexes from four dimensions: policy and law, market system, value creation and big data technology. The index system was scored by experts, and the entropy weight DEMATEL method was used to comprehensively evaluate the influencing factors of energy data market. The conclusion is that the key factors influencing the construction and development of energy data market are business compliance, data circulation system, product environmental benefit, product social benefit, data product development and data asset management. At present, the trading market of energy data is still in the primary stage, and the concrete implementation way of the construction and development of energy data market needs further study.
Reference


Proceedings of the 2\textsuperscript{nd} International Conference on Big Data Economy and Digital Management (BDEDM 2023)

January 6-8, 2023, Changsha, China

\textit{BDEDM 2023}

Copyright © 2023 EAI, European Alliance for Innovation

www.eai.eu

http://bdedm.net/

ISBN: 978-1-63190-402-8
European Alliance for Innovation

EAI is a non-profit organization with free membership and the largest open professional society for advancing research careers through community collaboration and fair recognition. Members benefit from finding feedback and mentorship for their work and they are guaranteed to be evaluated fairly, transparently, and objectively through community.

ISBN: 978-1-63190-402-8
ISSN: 2593-7642

http://eudl.eu/series/CORE | www.eai.eu