An Empirical Study on the Development Level of Sci-tech Finance for Anhui Province Based on Factor Model and Clustering Model

Zebin Liu[1], Guangkui Xu[2], Lei Shen[3]*, Yi Zhou[4]
zebin1992@163.com, 604692865@qq.com, *Corresponding author: Shenlei0922@qq.com, Zhouyi3280@163.com

School of Finance and Mathematics Huainan Normal University Huainan, P.R. China

Abstract: Sci-tech finance mainly refers to integrating the technology and financial industries, which belong to industrial finance. The world is seeing a considerably high degree of integration of technological industries in the era of the digital economy, which is beneficial for sci-tech finance to innovate and boost with the application of developing technologies like AI, cloud computing, quantum computing, blockchain, and big data. Sci-tech finance is instrumental in enhancing innovation and entrepreneurship, providing a new driving force for the tremendous enhancement of the regional economy in China. For the sake of quantifying the development level of sci-tech finance for Anhui province, this paper firstly calculates sci-tech finance indexes for 16 prefectural-level municipalities in Anhui province by using a factor model based on seven indicators covering three aspects: sci-tech finance input, sci-tech finance output and loan-level for financial institutions. Secondly, the clustering model analyzes each city's development level category based on a comprehensive index of sci-tech finance.

Keywords-component; Sci-tech finance, index calculation, factor model, cluster analysis

1 INTRODUCTION

The research from Changwen Zhao et al. [1] claimed that sci-tech finance is a set of financial instrument vehicles, financial systems, financial policies and financial services that promote the improvement of sci-tech, high-tech industries and the transformation of scientific and technological achievements, which plays an essential role in sci-tech innovation system and financial system. Zhao et al. constructed a general index system of sci-tech finance containing six-dimensional sub-indices: venture capital index, technology loan index, technology capital market index, technological insurance index, research and development index, and environment of sci-tech finance index. Ying Cao et al. [2] set up a comprehensive index of sci-tech finance for China, consisting of four sub-indices: sci-tech finance resource index, sci-tech finance funding index, sci-tech finance output and sci-tech finance loan index.

Jiang Wang [3] utilized the empirical method of combining panel data and the Malmquist index model for conducting a comparative study on the effectiveness of sci-tech finance for 30

MSIEID 2022, December 09-11, Chongqing, People's Republic of China

DOI 10.4108/eai.9-12-2022.2327633

provincial municipalities in China. Yulian Xu et al. [4] and Meng Zhang [5] used the provincial-level region in China as a decision-making unit to quantitively measure the allocative efficiency of capital for sci-tech finance by using the DEA-Malmquist model.

The main goal of this paper is to quantitively measure the development level of sci-tech finance on 16 prefectural-level municipalities in Anhui Province by calculating the comprehensive index of sci-tech finance, sci-tech finance input index, sci-tech finance output index and loan-size index for financial institutions. The framework of this paper is as follows. In Section 2, how the factor model calculates sci-tech finance indexes will be introduced. In section 3, this paper will discuss using cluster analysis to classify 16 prefectural-level municipalities based on the comprehensive index of sci-tech finance. The empirical analysis will be considered in Section 4. Conclusions will be drawn in Section 5.

2 CALCULATION OF SCI-TECH FINANCE INDEX

2.1 Construction of Sci-tech finance indicators

Considering the relevant research conducted by Zhao [1] and Cao [2], this paper constructs the indicator system for evaluating sci-tech finance for 16 prefectural-level municipalities in Anhui province. The indicator system is shown in Table 1. According to the feasibility and availability of data, data are mainly from the 2020 Statistical Yearbook of Anhui Province, the 2020 Statistical Bulletin of Anhui Province and the 2020 Statistical Yearbooks of prefectural-level municipalities in Anhui Province.

Table 1 Indicator system

Variables	**Sci-tech Finance Indicators**	
	Indicator Name	***Calculation Method***
X1	Human resources technology	Number of the active staff of technology/number of employed people
X2	R&D active staff	Number of R&D active staff/ number of employed people
X3	Level of financial support	Financial expenditure on science and technology/total financial expenditure
X4	R&D funding resources	R&D expenditure/ local GDP
X5	loan-level for financial institutions	Loan amount of financial institutions/local GDP
X6	Patent output rate	Number of patents granted/ R&D expenditure
X7	The turnover rate of the technical market	Turnover of technological market/ R&D expenditure

2.2 Construction of factor model

This paper adopts the factor model to calculate the sci-tech finance indexes, processing data in SPSS 26.0. Factor model is a statistical method for deterring common factors from a set of population data. It examines the internal dependencies among several variables and represents their underlying statistics structure with a few dummy variables.

The observable random vector *X* ($X=X_1, X_2..., X_p$) has *p* components, and the following formula is:

$$\begin{aligned} A_1 &= a_{11}F_1 + a_{12}F_2 + ... + a_{1m}F_m + \varepsilon_1 \\ A_2 &= a_{21}F_1 + a_{22}F_2 + ... + a_{2m}F_m + \varepsilon_2 \\ &\cdots \\ A_p &= a_{p1}F_1 + a_{p2}F_2 + ... + a_{pm}F_m + \varepsilon_p \end{aligned} \tag{1}$$

Or

$$\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pm} \end{bmatrix} \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_m \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_p \end{bmatrix} \tag{2}$$

F_1, $F_{2...}$, F_m are unobservable variables, called common factor, *p* additional sources of variation ε_i are the unique factor that the first m common factors cannot include. Matrix M=(a_{ij}) is called loading matrix, where a_{ij} is called loading that is the correlation coefficient of common factor F_m and variable X_p. It is crucial to observe which variables have larger loadings while obtaining common factors and then state the proper meaning of the common factor.

This paper employs the Z-score method to standardize the raw data into dimensionless values to remove the effect of different data outlines and quantitative levels on the empirical analysis. The calculation formula is as follows:

$$Z_{ij} = \frac{x_{ij} - \overline{x}_j}{s_i}, i = 1, 2, ..., n, j = 1, 2, ..., p \tag{3}$$

$$\overline{x}_j = \frac{\sum_{i=1}^{n} x_{ij}}{n}, s_j^2 = \frac{\sum_{i=1}^{n} (x_{ij} - \overline{x}_j)^2}{n-1} \tag{4}$$

2.3 Preliminary Analysis

This paper applies the method of Kaiser-Meyer-Olkin (KMO) and Bartlett's test to determine if the sample is adequate and appropriate for using the factor model. The KMO statistic e is assigned a value between 0 and 1. A zero value suggests that the sum of partial correlations is greater than the sum, implying that factor analysis is ineffective and vice versa. According to Kaiser, KMO values greater than 0.5 are acceptable. The null hypothesis, which is that the original correlation matrix is an identity, is tested by Bartlett's measure. The result shows that the KMO value is close to 0.7, and Bartlett's test is significant. The result is shown in table 2.

Table 2 KMO and BARTLETT'S Test

KMO		.698
Bartlett's Test	Approx. Chi-Square	68.615
	df	21
	Sig.	.000

2.4 Factor Extraction and Rotation

SPSS26.0 performs factor analysis on data from 16 cities in Anhui Province to obtain the total variance explained, and results are shown in Table 3 and Table 4. These two tables exhibit the extraction output, listing the eigenvalues related to each linear factor before and after rotation. According to the precept that the initial eigenvalue is greater than one, this paper extracts three common factors with approximately 89% of the cumulative variance contribution, and their eigenvalues are λ_1=3.761, λ_2=1.288 and λ_3=1.169 respectively, which indicates that three common factors can reflect most of the information of variables of the seven selected indicators. The rotation optimizes the factor structure, and one result for these data is that the relative significance of the three factors is equalized.

Table 3 Initial eigenvalues

	Initial Eigenvalues		
Component	**Total**	**% of Variance**	**Cumulative %**
1	3.761	53.724	53.724
2	1.288	18.403	72.128
3	1.169	16.698	88.826
4	.468	6.687	95.513
5	.112	1.596	97.109
6	.102	1.461	98.570
7	.100	1.430	100.000

Table 4 Rotation output

	Rotation Sums of Squared Loadings		
Component	**Total**	**% of Variance**	**Cumulative %**
1	3.632	51.885	51.885
2	1.313	18.756	70.641
3	1.273	18.184	88.826

The original loading matrix is constructed with common factors extracted by the principal component analysis method. For the sake of more reasonably explaining the connotations expressed by the common factors, the original component matrix is orthogonally rotated by Varimax with Kaiser Normalization [6]. Rotation can effectively maintain the orthogonality of each common factor, outputs as an exhibit in table 5. According to table 5, the first common factor F_1 consists of X_1, X_2, X_3 and X_4, which can be interpreted as sci-tech finance input resources reflecting investment in talent, finance and R&D institutions, and R&D funding in each city. The second common factor F_2 consists of X_6 and X_7, which can be interpreted as sci-tech finance output resources reflecting the output level in each city. The third common factor F_3 consists of X_5, which can be interpreted as loan size for financial institutions reflecting the degree of support for financial institutions in each city.

$$R_{VARIMAX=} \arg\max(\sum_{j=1}^{k}\sum_{i=1}^{p}(\Lambda R)_{ij}^{4} - \frac{\gamma}{p}\sum_{j=1}^{k}(\sum_{j=1}^{k}(\Lambda R)_{ij}^{2})^{2}) \tag{5}$$

Table 5 Rotated component matrix

Rotated Component Matrix			
	Component		
	F1	**F2**	**F3**
Zscore(X1)	.669	.080	.529
Zscore(X2)	.935	.156	.158
Zscore(X3)	.908	.251	.185
Zscore(X4)	.954	-.008	-.023
Zscore(X5)	.064	-.076	.926
Zscore(X6)	-.698	.616	.239
Zscore(X7)	.287	.913	-.137

2.5 Index Calculation

SPSS calculates the scores of each main factor as the index of sci-tech finance input (F_1), index of sci-tech finance output(F_2) and index of loan-size for financial institutions(F_3). The comprehensive index (F) of sci-tech finance is calculated as follow:

$$F = 0.584 \times F_1 + 0.211 \times F_2 + 0.205 \times F_3 \tag{6}$$

The weight of each common factor is the ratio of each factor's variance contribution after rotation to the total contribution. The comprehensive index is obtained by substituting each factor into formula (6), and the index results are exhibited in table 6.

Table 6 Index of sci-tech finance

City name	(F)	(F_1)	(F_2)	(F_3)
Hefei	1.86	1.99	1.41	1.98
Wuhu	1.04	1.80	0.49	-0.57
Ma Anshan	0.29	1.02	-1.11	-0.35
Bengbu	0.24	0.68	0.19	-0.95
Tongling	0.22	0.80	-1.24	0.07
Chuzhou	0.12	-0.39	1.95	-0.31
Xuancheng	-0.06	0.30	-0.59	-0.54
Huangshan	-0.27	-0.29	-0.96	0.47
Huaibei	-0.28	-0.14	-1.33	0.44
Huainan	-0.28	-0.92	1.04	0.18
Fuyang	-0.34	-1.19	0.23	1.48
Bozhou	-0.45	-1.04	0.69	0.05
Luan	-0.47	-0.95	-0.33	0.76
Anqing	-0.52	-0.41	-0.73	-0.60
Suzhou	-0.55	-0.85	-0.50	0.28
Chizhou	-0.56	-0.41	0.81	-2.38

3 CLUSTER ANALYSIS OUTPUT

Cluster analysis is a statistical method of classifying the objects studied by the data. This paper uses explicitly systematic cluster with an intergroup linkage approach using Euclidean distance to measure the average distance between samples based on the comprehensive index of sci-tech finance, mathematical expression as shown:

$$d_{xy} = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \tag{7}$$

where the clustering result are shown in figure 1.

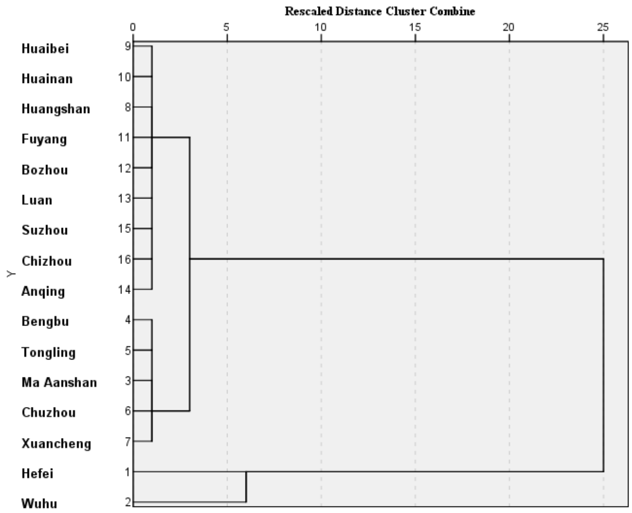

Figure 1. Clustering Output

4 EMPIRICAL ANALYSIS

According to the result of factor analysis, the clustering output of the development level of sci-tech finance for Anhui province can be divided into four categories. The first category is Hefei that is the city with the top development level of sci-tech finance. He Fei has the highest comprehensive index of 1.86, including that the index of loan-size for financial institutions and sci-tech finance input is ranked first, and the index of sci-tech finance output is ranked second. It has abundant sci-tech and financial resources and can utilize various resources to develop sci-tech finance. According to the ranking of research cities published in the UK's Nature supplement, it ranked third in China in terms of urban research strength in 2019. It has become the most potent area of effective integration of sci-tech finance in Anhui province.

The second category is Wuhu, the city with a higher development level of sci-tech finance, and it has the second-highest comprehensive index of sci-tech finance. It has a particular gap between the development level of sci-tech finance with Hefei, but it also has a good infrastructure and ample funds for developing sci-tech finance. Recently, it has deeply implemented the innovation-driven development strategy, fully integrated into the Yangtze River Delta Sci-tech Innovation Community and G60 Sci-tech Innovation Corridor, and continuously optimized the innovation policy environment. It has great potential for developing sci-tech finance to be a competitive "rival" of Hefei in the future.

The third category is Ma Anshan, Bengbu, Tongling, Chuzhou, and Xuancheng, the city with a medium-level development of sci-tech finance. The index of sci-tech finance input, sci-tech finance output and loan-size for financial institutions show the problem of the unbalanced development giving rise to the comprehensive index of sci-tech finance lower than Hefei and

Wuhu. These cities have a certain foundation for developing sci-tech finance, and they have considerable room for future improvement.

The fourth category is Huangshan, Huaibei, …, and Chizhou, the city with a low development level of sci-tech finance. These cities have obvious deficiencies in resources associated with sci-tech finance, which restricts the development of sci-tech finance because the sci-tech and finance cannot be highly integrated.

5 CONCLUSIONS

According to the outputs calculated by factor analysis, Hefei has the highest comprehensive index of 1.86, and Chizhou has the lowest comprehensive index of-0.56. In terms of comprehensive index, there are six cities with a positive index and ten cities with a negative index. The present work shows that the development level of sci-tech finance in Anhui province is unbalanced, and there are considerable differences between these cities.

Based on the result of clustering analysis, these 16 cities can be divided into four categories based on their comprehensive index of sci-tech finance: top-level, higher-level, medium-level, and low-level. Hefei and Wu Hu have an advanced development level of sci-tech finance, while other cities have obvious gaps with these two cities, in which the contribution of sci-tech input resources is the largest; nevertheless, the contribution of output level and loan-size level is relatively tiny.

Acknowledgment. The authors gratefully acknowledge the support of Natural Science Research Projects in Anhui Universities (KJ2020A0643), Huainan Normal University Research Projects (2021XJYB005), (2021XJZD007) and (2021XJZD004)

REFERENCES

[1] C.W.Zhao,(2009) Sci-tech Finance.Science Press, Beijing.

[2] J.B.Liu and HY.Ma, "Evaluation of the combination of sci-tech finance in strategic emerging industries",Science and Technology Management Research,vol.2014,no.15,pp.115-119,2014.

[3] J.Wang, "Sci-tech progress environment, Hi-tech industrialization, sci-tech financial benefit and transformation of economic growth mode", Science and Technology Management Research, vol.2015, no.14, pp.32-37, 2015.

[4] Y.L.Xu and Y.D.Wang, "Launch of the GEM, corporate financing constraints and R&D intensity", Soft Science, vol.2015, no.8, pp.53-56, 2015.

[5] M.Zhang, "Research on the operational efficiency of sci-tech finance in Shandong Province", Shandong University of Finance and Economics, vol.2018, 2018

[6] H.F.Kaiser, "The varimax criterion for analytic rotation in factor analysis", Psychometrika, vol.23, no.3,1958.