Application of Factor Analysis and Cluster Analysis in Regional Economics
A Case Study of Beijing Tianjin Hebei Urban Agglomeration

Jia Li
e-mail: 20120524@bjtu.edu.cn

Beijing Jiaotong University Beijing, China

Abstract—This paper uses cluster analysis and factor analysis to study the economic status of Beijing Tianjin Hebei Urban Agglomeration. We include the variables such as the regional GDP and the general public budget income of 167 districts and counties in the urban agglomeration into the public factors, then use the public factors and their factor contribution rate to construct the comprehensive factor score, then conduct cluster analysis according to the comprehensive factor score ranking, and finally conduct comparative analysis.

Keywords-cluster analysis; factor analysis; Beijing-Tianjin-Hebei Urban Agglomeration

1 INTRODUCTION

Cluster analysis is a multivariate statistical analysis method. The application of this method is to classify a large number of sample data according to their attributes without any prior knowledge and reference mode. It makes individuals in the same class have greater similarities and individuals in different classes have greater differences. In essence, this is a classification method in which birds of a feather flock together. A dimension reduction method often used with it is factor analysis. Factor analysis is a statistical technique for extracting common factors from variable groups, which was first proposed by Spearman. Its main purpose is to describe some more basic hidden variables hidden in a group of measured variables, but they can not be measured directly.

The combination of cluster analysis and factor analysis is very common. For example, G. Q. Zhou and Y. R. Li found that the performance evaluation combined with factor analysis and cluster analysis can well evaluate the road traffic poverty alleviation [1], while X. Zhao used this method to study the economic situation of Shandong Province. This paper will use this method to study the related problems of the urban agglomeration economy [2].

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

DOI 10.4108/eai.9-12-2022.2327702

Urban agglomeration plays a huge role in driving economic development. In August 2017, the China Development Research Foundation team released its findings: In 2015, 12 urban agglomerations, including Yangtze River Delta, Pearl River Delta, Beijing-Tianjin-Hebei (BTH) region, the Western Coast of Taiwan Strait, Shandong Peninsula, Central Plains, Wuhan, Changsha-Zhuzhou-Xiangtan region, Guanzhong plain, Chengdu-Chongqing region, the central-southern Liaoning and Harbin-Changchun region, gathered 80% of China's total economic output. The total area of these 12 urban agglomerations accounts for less than 20% of China's land area and more than 60% of its population. Therefore, the state attaches great importance to the top-level design of urban group development. On November 18, 2018, the Opinions of the CPC Central Committee and the State Council on the Establishment of a More Effective New Mechanism for Regional Coordinated Development issued by the CPC Central Committee and the State Council clearly stated that the national major regional strategic integration and development should be promoted by cities such as BTH Urban Agglomeration, Yangtze River Delta Urban Agglomeration, Guangdong-Hong Kong-Macao Greater Bay, Chengdu-Chongqing Urban Agglomeration, Yangtze River Midstream Urban Agglomeration, Central Plains Urban Agglomeration and Guanzhong Plain Urban Agglomeration. Among them, the coordinated development of BTH Urban Agglomeration is a major national strategy personally planned and personally promoted by China's President Xi Jinping, the core of which is to solve the functions of Beijing's non-capital in an orderly manner, adjust the economic structure and spatial structure, walk out of a new way of intensive development, explore a model of optimizing development in densely populated areas, promote coordinated regional development and form a new growth pole. Therefore, this paper will focus on the economic development status of BTH Urban Agglomeration.

BTH Urban Agglomeration is located in the North China Plain and adjacent to the Bohai Bay in the East. It is the core area of northern China's economy. The integrated development of BTH Urban Agglomeration is different from that of Yangtze River Delta, Pearl River Delta and other urban agglomerations. There is an obvious administrative separation between Beijing, Tianjin and Hebei. There are a number of "broken roads" such as Beijing Qinhuangdao Expressway and Beijing Taiwan Expressway at the provincial boundary. The factor mobility among the three places is very poor, and the economic development of the three places is uneven. In real life, it can be seen that more enterprises prefer to squeeze into the capital with high land price, nor do they want to be located in Hebei, which is adjacent to it. Therefore, the development of BTH Urban Agglomeration depends more on policy guidance. The proposal of coordinated development of Beijing, Tianjin and Hebei is to break the barriers of factor market, promote resource sharing and complementarity among the three places, and stimulate market and social vitality. In the coordinated development plan of Beijing, Tianjin and Hebei, the positioning of the three provinces and cities is different: Beijing's goal is to become a "national political center, cultural center, international communication center and scientific and technological innovation center"; Tianjin is positioned as "national advanced manufacturing R & D base, north international shipping core area, financial innovation and operation demonstration area and reform and opening up pilot area"; Hebei Province will develop into a "national important base for modern trade and logistics, an experimental area for industrial transformation and upgrading, a demonstration area for new urbanization and urban-rural overall planning, and an ecological environment support area for Beijing-Tianjin-Hebei Urban Agglomeration ".

It has been seven years since the BTH coordinated development strategy was put forward in 2014. Although there is still a big gap in the economic development of the three places, J. Q. Yuan and others believed that the coordinated development strategy of BTH has promoted the economic development of the three places, especially in areas with low per capita GDP [3]. P. Zeng evaluated and compared the comprehensive development level of China's top ten urban agglomerations. According to the comprehensive development ability, the top ten urban agglomerations are divided into three categories: strong, relatively general and relatively low. Among them, BTH ranks second, belonging to zone I with strong comprehensive development ability [4]. However, L. Y. Wei believed that the coordinated development policy of BTH has restrained the economic development of Beijing and Tianjin to the greatest extent, At the same time, it has also accelerated the pace of economic development in Hebei Province to the greatest extent [5]. F. Q. Niu found that the central cities of BTH developed under the policy guidance of the 1990s, and the tertiary industry developed rapidly, but its manufacturing industry did not develop fully [6]. J. Luan and R. Ma believed that from 2000 to 2018, the trend of unbalanced internal development in BTH is more significant [7]. The influence of existing literature on the coordinated development of BTH is controversial. The innovation of this paper is that it adopts the data at the county level. It is more practical to investigate the driving role of core cities in the surrounding areas with smaller administrative units. It can also be used to observe the implementation of collaborative planning spatial layout and check the specific completion of the BTH coordinated development strategy at the county level in the seven years of planning.

2 MATERIALS AND METHODS

The data of this paper comes from the Guotai'an database, China’s economic and social research big data platform, the EPS data platform and yearbooks of some provinces and cities. Due to the availability of data, this paper selects the data of 2010 and 2018. The research methods used in this paper are as follows.

2.1 Factor Analysis

Based on the principal component analysis, this paper integrates the variables such as the regional GDP, the output value of primary, secondary and tertiary industries, the general public budget revenue and the general public budget expenditure into public factors, in which the missing value of the original variables is filled with the average value. Then, by calculating the factor score, evaluate the performance of the economic development of the three districts and counties in 2010 and 2018.

2.2 Cluster Analysis

Using several selected public factors as clustering variables in cluster analysis, cluster analysis was carried out for each district and county in 2010 and 2018 according to different years, and the final results were presented in tables. In this paper, the method of calculating the distance between classes is the sum of squares of deviations.

2.3 Comparative Analysis

Based on factor analysis and cluster analysis, this paper shows the changes of public factor scores and levels of economic development of districts and counties in BTH in 2010 and 2018 in the form of tables, compares the improvement or decrease of scores and the levels of districts and counties, and then combines their geographical location Policy orientation and other specific circumstances, make a specific analysis of the score change and grade change, and summarize the reasons.

3 RESULTS

In this paper, two common factors are extracted by the principal component analysis (See Table 1 for details). After rotating the component matrix, we can see the original indicators in each common factor. As can be seen from Table 2, public factor 1 (X_1) mainly covers the indicators of the overall strength of the regional economy and the strength of the local government, such as the regional GDP, the general public budget revenue, the general public budget expenditure, the added value of the tertiary industry; Public factor 2 (X_2) includes indicators reflecting the basic situation of local industry and agriculture, such as the added value of public primary and secondary industries and the number of industrial enterprises above designated size.

Table 1 Factor contribution rate table

Component	**Initial eigenvalue**			**Extract the sum of squares of loads after rotation**		
	total	***Percentage variance***	***Cumulat-ive%***	***total***	***Percentage variance***	***Cumul-ative%***
1	4.521	64.579	64.579	4.177	59.665	59.665
2	1.335	19.066	83.645	1.679	23.981	83.645

Table 2 Rotating component matrix

Initial Variable	***X1***	***X2***
Regional GDP	0.962	0.182
Added value of primary industry	-0.421	0.574
Added value of secondary industry	0.527	0.739
General public budget revenue	0.967	0.155
General public budget expenditure	0.915	0.243
Number of Industrial Enterprises above Designated Size	0.305	0.827
Added value of tertiary industry	0.965	0.043

Then, this paper constructs the comprehensive factor score $S_{i,j}$ through the above common factors X_1 and X_2 and their factor contribution rates a_1 and a_2. See (1) for the specific formula. i represents the year and j represents the district and county. Then, the economic development ranking of districts and counties in BTH in 2010 and 2018 is calculated according to the comprehensive factor score.

$$S_{i,j}= a_1*X_{j,1}+a_2* X_{j,2} \quad (1)$$

The ranking in 2018 minus the ranking in 2010 gives us Table 3. This paper intercepts the ranking changes of the top 30 districts and counties in 2018.

Table 3 Ranking changes of the top 30 districts and counties in 2018

Ranking in 2018	Name of districts or counties	Change of ranking	Ranking in 2018	Name of districts or counties	Change of ranking
1	Haidian	0	16	Qian'an	4
2	Chaoyang	0	17	Dongli	-7
3	Xicheng	0	18	Gaocheng	78
4	Shunyi	0	19	Fengrun	-2
5	Wuqing	11	20	Baodi	6
6	Dongcheng	6	21	Fengnan	-3
7	Xiqing	-2	22	Renqiu	101
8	Tongzhou	-1	23	Wu'an	74
9	Fengtai	4	24	Caofeidian	10
10	Beichen	-1	25	Xinji	29
11	Daxing	3	26	Shijingshan	11
12	Changping	-4	27	Huairou	11
13	Jinghai	11	28	Sanhe	111
14	Fangshan	-8	29	Hexi	3
15	Jinnan	0	30	Miyun	5

Comparing the ranking changes of the comprehensive factor scores of districts and counties in Beijing, Tianjin and Hebei in 2010 and 2018, it can be found that the ranking changes of the previous top districts and counties are small, among which the rankings of Wuqing and Jinghai in Tianjin have increased, while the rankings of some districts in Beijing have declined, such as Fangshan, Changping and other outer suburbs. The rankings of Gaocheng, Qian'an, Caofeidian, Sanhe, Renqiu and Zhuozhou in Hebei Province have improved significantly. Compared with Zhangjiakou, Chengde and other urban areas, the geographical locations of these districts and counties in Hebei are more southern. This may be because the areas covered by Beijing, Tianjin and Shijiazhuang, the capital of Hebei Province, in the BTH coordinated development plan are in the south of the BTH Urban Agglomeration. At the same time, most districts and counties in

Langfang, Baodi, Cangzhou, Tangshan and other cities in Hebei have improved their ranking. To a certain extent, it shows the driving effect of the coordinated development of Beijing, Tianjin and Hebei on the surrounding areas, and it also shows that this driving effect needs to be further improved. The decline in the ranking of Beijing districts and counties and the rise in the ranking of Tianjin and Hebei districts and counties confirm that the internal difference of BTH Urban Agglomeration decreases.

However, despite the changes in the ranking of districts and counties in 2010 and 2018, Haidian, Chaoyang, Xicheng and Shunyi in Beijing are still in the top four. Among the top 20, Beijing and Tianjin still account for the majority. Therefore, there is still a gap in the economic status among cities in BTH Urban Agglomeration.

Then, according to the above comprehensive factor scores, 167 districts and counties are divided into four levels and three levels according to different years by cluster analysis. The results for 2010 and 2018 are shown in Tables 4 and Table 5.

Table 4 Districts and counties level division of BTH Urban Agglomeration in 2010

Rank	**Name of districts or counties**
First level	Haidian, Chaoyang
Second level	Xicheng, Shunyi
Third level	Xiqing, Fangshan, Tongzhou, Changping, Beichen, Dongli, Qian'an, Dongcheng
Fourth level	Other regions

Table 5 Districts and counties level division of BTH Urban Agglomeration in 2018

Rank	**Name of districts or counties**
First level	Haidian, Chaoyang, Xicheng
Second level	Shunyi, Wuqing, Dongcheng, Xiqing, Tongzhou, Fengtai
Third level	Other regions

Through comparative analysis, we found that the number of levels of districts and counties in BTH Urban Agglomeration was three in 2018, down from four in 2010. We believe that the vertical gap between districts and counties in BTH Urban Agglomeration is gradually narrowing. Among the variables selected in this paper, there are many industrial measurement indicators. Therefore, it can be inferred that the difference between districts and counties within the BTH Urban Agglomeration is reduced, and the industrial transfer plays a certain role in promoting the development of Hebei and Tianjin. For example, the establishment of Wuqing-Beijing-Tianjin industrial new city and Caofeidian circular economy demonstration zone has promoted the growth of the local economy.

While the level is flatter, Beijing Haidian and Chaoyang are still in the leading position. In the district and county level division of BTH Urban Agglomeration in 2018, there are no districts and counties in Hebei Province in the first and second levels, indicating that the economic development potential of Hebei Province still needs to be further tapped, but the differences within the city are narrowing.

4 CONCLUSIONS

With the proposal of the coordinated development strategy of BTH, more and more scholars put forward their thoughts on the balance of the development of BTH Urban Agglomeration. There is a debate in the existing literature on the impact of the coordinated development of BTH on the three places, and most of the current research measures the economic indicators at the urban level or the industrial level, and there is little research specific to the county area. Therefore, this paper intends to use county data to further study the development status of BTH Urban Agglomeration based on factor analysis and cluster analysis.

This paper finds that there is a gap in economic status between cities in BTH Urban Agglomeration, but the difference within the urban agglomeration decreases. At the same time, this paper draws the following conclusions: (1) the areas covered by Beijing, Tianjin and Shijiazhuang, the capital of Hebei Province, are in the south of BTH Urban Agglomeration, while the comprehensive factor score ranking of districts and counties in Hebei Province located in this area has improved significantly after the implementation of the Beijing Tianjin Hebei coordinated development plan; (2) In the level division of districts and counties in BTH Urban Agglomeration in 2018, there are no districts and counties in Hebei Province in the first and second levels, indicating that the economic development potential of Hebei Province still needs to be further tapped.

Based on the above conclusions, the enlightenments significance of this study to the development of BTH Urban Agglomeration are as follows.

4.1 The driving role of the coordinated development of BTH still needs to be further improved.

Haidian and Chaoyang in Beijing have been in the leading position in 2010 and 2018. Among the top 20 cities, Beijing and Tianjin still account for the majority. Therefore, there is still a gap in the economic status between the cities of BTH Urban Agglomeration, indicating that the economic development potential of Hebei Province still needs to be further tapped, and the driving role of Beijing and Tianjin core cities still needs to be further improved.

4.2 Promote the construction of the northern part of the urban agglomeration.

The southern districts and counties of the BTH Urban Agglomeration have received some help in the coordinated development of BTH. Compared with the situation before the plan, the economy of these areas has been significantly improved, but the northern districts and counties still have great development potential. Therefore, the 2022 Winter Olympics should be held to promote the development of tourism in the north of BTH Urban Agglomeration. Give full play

to the leading role of the preparation for the 2022 Winter Olympic Games, vigorously develop sports, culture, tourism and leisure, convention and exhibition and other eco-friendly industries in Zhangjiakou and Chengde, and jointly build the Beijing-Zhangjiakou culture and sports tourism belt.

REFERENCES

[1] G. Q. Zhou and Y. R. Li, “Performance evaluation of road traffic poverty alleviation based on factor cluster analysis,” in Journal of Railway Science and Engineering, vol. 18, X. D. Jiang, Eds. 2021, pp. 2490-2496.

[2] X. Zhao, “Study on the economic situation of Shandong Province Based on principal component cluster analysis,” in China's Collective Economy, vol. 25. 2021, pp. 26-27.

[3] J. Q. Yuan, W. Bu, and Y. X. Yang, “How to break through the ‘double low-end locking’ of Beijing, Tianjin and Hebei -- Research on industrial upgrading and economic growth effect based on the regional value chain,” in Industrial Economic Research, vol. 5, F. Y. Dai, Eds. 2019, pp. 13-26.

[4] P. Zeng, “Comprehensive development level of China's top ten urban agglomerations: factor analysis and comprehensive integrated evaluation,” in China Population, Resources and Environment, vol. 1, X. J. Wang, Eds. 2008, pp. 69-73.

[5] L. Y. Wei, “Investigation on the macroeconomic effects of the coordinated development policy of Beijing, Tianjin and Hebei,” in Statistics and Decision Making, vol. 35, L. Q. Liu, Eds. 2019, pp. 123-127.

[6] F. Q. Niu, X. Y. Yang, and F. Wang, “Urban Agglomeration Formation and Its Spatiotemporal Expansion Process in China: From the Perspective of Industrial Evolution,” in Chinese Geographical Science, vol. 30. 2020, pp. 532-543.

[7] J. Luan and R. Ma, “Statistical measurement of economic coordinated development in Beijing Tianjin Hebei region,” in Statistics and decision making, vol. 36, L. Q. Liu, Eds. 2020, pp. 50-54.