Competitiveness evaluation and cluster analysis of logistics industry in Henan Province based on SPSS 22.0

Qianqian Liu

1College of Transport & Communications, Shanghai Maritime University, Shanghai, 201306, China

*Corresponding author’s e-mail: 284732060@qq.com

Abstract. The logistics industry is an important support for the construction and development of my country's national economy, and there is a close relationship between regional logistics and regional economy. In order to comprehensively evaluate the level of logistics competitiveness in Henan Province, this paper has established 11 evaluation indexes of Henan logistics industry competitiveness from three perspectives: regional economic conditions, logistics demand and logistics industry support factors. With the help of SPSS 22.0 software for factor analysis, 11 indexes are divided into two public factors: logistics economic development environment, logistics demand and scale, and the competitiveness of the logistics industry of 18 province-administered cities in Henan Province is ranked. According to the scores of 2 factors, the 18 provincial cities were divided into five categories using the systematic cluster analysis method, and then the development ideas for the development of the logistics industry of each type of provincial cities were put forward to provide a reference for the overall development of the logistics industry in Henan Province. At the same time, it also provides reference for the development of logistics industry in other regions.

1. Introduction

With the transformation of my country's economy to high-quality development, the logistics industry has gradually developed towards intensive, intelligent and standardized development. But for a long time, my country's logistics cost is high, which directly affects the operation level of the national economy. Therefore, reducing costs and increasing efficiency and high-quality development have become the long-term goals pursued by the logistics industry. Since the 18th National Congress of the Communist Party of China, General Secretary Xi Jinping has given a series of important instructions to the logistics industry, pointing out the direction for the logistics industry to reduce costs and increase efficiency. The state has also issued a number of policy opinions one after another, proposing precise measures to reduce costs and increase efficiency in the logistics industry. In recent years, the ratio of China's total social logistics costs to GDP has been declining year by year, and the logistics industry has achieved substantial results in "cost reduction".

In recent years, China's logistics industry has developed rapidly. Many regions have taken the development of modern logistics industry as a pillar industry, and have adopted various measures to enhance the competitiveness of the logistics industry, and promote healthy and sustained economic growth through the development of logistics. Henan is located in the central area of my country, connecting the three major economic regions of the east, middle and west, and has the regional advantage of developing the logistics industry. However, in fact, the development level of the overall logistics industry and the economic value generated in Henan Province does not have an advantage in
the national ranking. The development of regional logistics is closely related to the development of the regional economy. Therefore, how to realize the leap-forward development of the regional logistics industry is a problem that must be considered in accelerating the development of Henan's economic construction. Whereas, constructive ideas put forward in the previous logistics development, the paper using a more sophisticated factor analysis of the main Henan 18 cities in the competitiveness of the logistics for scientific evaluation and research, and then to develop a reasonable regional logistics development strategy.

2. Construction of evaluation index system
Although many scholars have done research on the competitiveness of logistics industry in our country or some provinces, there is still no perfect evaluation system for the evaluation index of regional logistics competitiveness. The selected indicators should be able to facilitate the comparative analysis of the logistics development of each city, and the selection of the logistics development capability evaluation indicators must be objective and reasonable, so as to ensure the scientific construction of the evaluation system.

In this paper, considering the three perspectives of regional socio-economic conditions, logistics demand and logistics industry support factors, 18 cities in Henan Province are selected as samples to study the competitiveness of Henan Province's logistics industry.

It is planned to select the following 11 index variables to establish an evaluation index system for the competitiveness of the logistics industry in Henan Province for empirical research as shown in Table 1. Regional economic conditions include regional GDP (X1), per capita GDP (X2), permanent population (X3); logistics needs include the proportion of fixed asset investment in transportation, storage and postal industry (x4), total retail sales of consumer goods (x5), freight volume (x6), cargo turnover (x7); supporting elements of the logistics industry include highway mileage (x8), number of trucks (x9), logistics industry employees (X10), and the number of industrial enterprises above designated size (X11).

| Evaluation on the Competitiveness of Henan Logistics Industry | Number of industrial enterprises above designated size | GDP X1 | GDP per capita X2 | Permanent population X3 | Proportion of fixed asset investment in transportation, warehousing and postal industry X4 | The total retail sales of social consumer goods X5 | Freight volume X6 | Cargo turnover X7 | Highway route mileage X8 | Number of trucks X9 | Practitioners in the logistics industry X10 | Number of industrial enterprises above designated size X11 |
|---------------------------------------------------------------|------------------------------------------------------|--------|------------------|------------------------|----------------------------------------------------------------|----------------------|------------------|------------------|------------------------|------------------|--------------------------------------------|-----------------------------|
| Logistics demand                                              |                                                      |        |                  |                        | (1) Regional economic conditions include regional gross product (X1), per capita GDP (X2), permanent population (X3). | (2) Logistics demand includes the proportion of fixed asset investment in transportation, warehousing and postal industry (x4), total retail sales of consumer goods (x5), freight volume (x6), and cargo turnover (x7). |
(3) The supporting elements of the logistics industry include highway mileage (x8), the number of trucks (x9), logistics industry employees (X11), and the number of industrial enterprises above designated size (X12).

3. An empirical analysis of the competitiveness of the logistics industry in Henan province

3.1. Data source
Since COVID-19 broke out worldwide in year 2020, data of year 2020 cannot accurately reflect the actual level of Henan Province. Therefore, the year 2019 is chose as the original data, the establishment of appropriate statistical indicator system, all data are from "2019 Henan Statistical Yearbook", the sample data used in this article are all from the statistical yearbook published on the official website of the Henan Provincial Bureau of Statistics in 2019.

3.2. Factor analysis
In the process of empirical analysis of the competitiveness of the logistics industry in Henan Province, in order to eliminate the difference in data dimensions, first of all, the sample data of the 18 cities in Henan shown in the above table should be standardized using Spss22.0 software. Process, and then perform KMO and Bartlett sphere tests based on the standardized sample data to determine whether it is suitable for factor analysis. Therefore, in the process of inputting the sample data into the Spss22.0 software for analysis, through the inspection shown in Table 2, we found that the KMO value is 0.660>0.5 and the approximate chi-square value of the Bartlett sphere test is 321, and the degree of freedom is 55. The significance probability P value of the test is 0.000, indicating that it is suitable for factor analysis.

| Table 2. KMO and Bartlett sphere test results |
|---------------------------------------------|
| KMO and Bartlett test                        |
| Kaiser-Meyer-Olkin measurement taken sample adequacy. | .660 |
| Bartlett’s sphere test                       | Approximately Chi-square | 321.386 |
|                                            | df                        | 55      |
|                                            | Significance               | .000    |

According to the standardized sample data, the correlation coefficient matrix is obtained, and the common factor variance table can be obtained by using Spss22.0 software. The common factor variance indicates the degree of explanation of each variable by the extracted common factor. The common factor variance table of this sample data, the commonality of most variables is between 0.7-0.9, indicating that the common factor can explain the information of each variable to a higher degree, indicating the selected indicators and the collected data obtained can well reflect the level of competitiveness of the logistics industry in 18 cities in Henan Province, and provide basic support for the extraction of common factors below.

Then, according to the standardized correlation coefficient matrix, the eigenvalues of the correlation coefficient matrix and the cumulative variance contribution rate are calculated, and the number of common factors that should be selected is judged. In this paper, using Spss22.0 software, the standardized correlation coefficient matrix can be used to obtain the explained total variance table. As shown in Table 4, we can find that after the maximum variance orthogonal rotation, the first two eigenvalues are all greater than 1, and the cumulative contribution of the first two components reaches 85.764%, which is greater than 80%, which can be effectively retained the original data covers the original data information. Therefore, it is determined to select two principal components for factor analysis to complete the research on the regional logistics industry competitiveness evaluation.
Table 3. shows the statistics of the number of variables

| element | starting eigenvalue | retrieve the sum of squares | cyclic sum of squares loading |
|---------|---------------------|----------------------------|--------------------------------|
|         | total               | mutate                      | accumul %                      | total           | mutate %                    | accumul %                      |
| 1       | 7.339               | 66.714                      | 66.714                         | 7.339           | 66.714                      | 66.714                         |
| 2       | 2.095               | 19.049                      | 85.764                         | 2.095           | 19.049                      | 85.764                         |
| 3       | .765                | 6.951                       | 92.715                         | .765            | 6.951                       | 92.715                         |
| 4       | .428                | 3.889                       | 96.604                         | .428            | 3.889                       | 96.604                         |
| 5       | .177                | 1.608                       | 98.212                         | .177            | 1.608                       | 98.212                         |
| 6       | .097                | .879                        | 99.092                         | .097            | .879                        | 99.092                         |
| 7       | .068                | .619                        | 99.710                         | .068            | .619                        | 99.710                         |
| 8       | .015                | .137                        | 99.847                         | .015            | .137                        | 99.847                         |
| 9       | .011                | .101                        | 99.948                         | .011            | .101                        | 99.948                         |
| 10      | .004                | .035                        | 99.984                         | .004            | .035                        | 99.984                         |
| 11      | .002                | .016                        | 100.000                        | .002            | .016                        | 100.000                        |

Acquisition method: analysis of main components.

Through the explained total variance table, the cumulative contribution rate of the two principal components obtained reached 85.764%, indicating that the selected two principal components can reflect most of the basic indicator information. Then use the Spss22.0 software to obtain the orthogonal rotation component matrix shown in Table 6 according to the orthogonal rotation, and obtain the more distinct common factor components, and complete the basic explanation of the specific indicators.

Table 4. Rotation component matrix

| Rotating element matrix a |
|---------------------------|
| element                   | 1      | 2      |
| Zscore (Gross Regional Product) | .909   | .374   |
| Zscore (GDP per capita)    | .126   | .940   |
| Zscore (Resident population) | .914   | -.271  |
| Zscore (Proportion of investment in fixed assets of transportation and warehousing) | -.154  | .752   |
| Zscore (Total retail sales of consumer goods) | .945   | .273   |
| Zscore (Shipping volume)   | .875   | -.109  |
| Zscore (Cargo turnover)    | .855   | -.167  |
| Zscore (Highway route mileage) | .848   | .415   |
| Zscore (Number of trucks)  | .944   | -.098  |
| Zscore (Practitioners in the logistics industry) | .790   | .515   |
| Zscore (Number of industrial enterprises above designated size) | .941   | .037   |

Acquisition method: analysis of main components.
Rotation method: Maximum variation method with Kaiser normalization.
a. Convergence loops in 3 iterations.
It can be seen from Table 4 that the principal component 1 (Y1) is mainly composed of the variables X1 regional GDP, X3 permanent population, X5 total retail sales of consumer goods, X6 freight volume, X7 cargo turnover, X8 highway route mileage, X9 cargo determined by the number of automobiles, employees in the logistics industry in X10 countries, and employees in industries above the designated size of X11, their factor loads are 0.909, 0.914, 0.945, 0.875, 0.855, 0.848, 0.944, 0.790, and 0.941, respectively, reflecting the social economy at the field level, the environment, logistics operation status, and related industry development level indicators are a collection of regional logistics environmental competitiveness and potential competitiveness, which can be named as the economic and social development factor of the regional logistics industry. Since the cumulative variance contribution rate of Y1 has reached 65.537%, so it is an aspect that needs to be considered when studying the competitiveness of regional logistics.

It can be seen from Table 4 that the principal component 2 (Y2) is mainly determined by the variables D2 per capita GDP and the proportion of D4 transportation and storage fixed assets investment. Their factor loads are 0.940 and 0.752 respectively. The per capita GDP load coefficients of the two main factors are relatively high. Large, named as the logistics demand factor. The cumulative variance contribution rate of Y2 is 20.227%, which is also an important aspect that needs to be considered when studying regional logistics competitiveness.

Using Spss22.0 software, using regression analysis method, according to the proportion of the eigenvalue of the main factor as the weight, the factor score coefficient matrix is obtained, as shown in Table 5.

| Rating system element number matrix |
|-------------------------------------|
| element | 1 | 2 |
| Zscore (Gross Regional Product) | .112 | .127 |
| Zscore (GDP per capita) | -.031 | .434 |
| Zscore (resident population) | .147 | -.176 |
| Zscore (Proportion of investment in fixed assets of transportation and warehousing) | -.062 | .361 |
| Zscore (total retail sales of consumer goods) | .122 | .078 |
| Zscore (shipping volume) | .132 | -.098 |
| Zscore (cargo turnover) | .133 | -.124 |
| Zscore (Highway route mileage) | .101 | .150 |
| Zscore (number of trucks) | .142 | -.096 |
| Zscore (Practitioners in the logistics industry) | .087 | .200 |
| Zscore (number of industrial enterprises above designated size) | .134 | -.033 |

Acquisition method: analysis of main components.  
Rotation method: Maximum variation method with Kaiser normalization.  
Component score.

According to Table 5, the following functions can be obtained:

\[ Y1 = 0.112 \times X1 - 0.031 \times X2 + 0.147 \times X3 - 0.062 \times X4 + 0.122 \times X5 + 0.132 \times X6 + 0.133 \times X7 + 0.101 \times X8 + 0.142 \times X9 + 0.087 \times X10 + 0.134 \times X11 \]
\[ Y_2 = 0.127X_1 + 0.434X_2 - 0.176X_3 + 0.361X_4 + 0.078X_5 - 0.124X_7 + 0.150X_8 - 0.096X_9 + 0.200X_{10} - 0.033X_{11} \]

According to the explained total variance table (Table 4), the variance contribution rate after the rotation of the two common factors is used as the weight, and the score of each factor is weighted to obtain the comprehensive score. The comprehensive score function is as follows:

\[ Y = 65.537\%Y_1 + 20.227\%Y_2 \]

Then bring the collected sample data into the above-mentioned factor score function and comprehensive score function, and calculate the comprehensive score of 18 cities in Henan Province and the scores of the specific two factors. As shown in Table 6, the comprehensive ranking of 18 cities in Henan Province is as follows:

| Sample area | overall ratings  | total ranking | Y1 score | Y1 ranking | Y2 score | Y2 ranking |
|-------------|------------------|---------------|----------|------------|----------|------------|
| Zhengzhou   | 2.35             | 1             | 2.87006  | 1          | 2.32216  | 2          |
| Kaifeng     | .51              | 3             | .99700   | 3          | -.06452  | 7          |
| Luoyang     | .64              | 2             | 1.03042  | 2          | -.82902  | 15         |
| Pingdingshan| .20              | 4             | .67214   | 4          | 1.19806  | 18         |
| Anyang      | .15              | 5             | .56471   | 5          | 1.09812  | 17         |
| Hebi        | -0.02            | 6             | -2.7945  | 12         | .56008   | 4          |
| Xinxiang    | -0.06            | 7             | -2.4612  | 10         | .33085   | 5          |
| Jiaozuo     | -.14             | 10            | .11775   | 7          | -6.7808  | 14         |
| Puyang      | -.18             | 11            | .23265   | 6          | -8.6854  | 16         |
| Xuchang     | -.09             | 9             | -.09377  | 8          | -.39380  | 12         |
| Luoye       | -.21             | 12            | -.21078  | 9          | -.22163  | 11         |
| Sanmenxia   | -.07             | 8             | -.27720  | 11         | -.13310  | 9          |
| Nanyang     | -.40             | 14            | -.58112  | 18         | 2.36460  | 1          |
| Shangqiu    | -.38             | 13            | -.61153  | 14         | .08396   | 6          |
| Xinyang     | -.58             | 17            | -.93584  | 16         | .67151   | 3          |
| Zhoukou     | -.48             | 15            | -.42754  | 13         | -.57489  | 13         |
| Zhumadian   | -.67             | 18            | -.84604  | 15         | -.12751  | 8          |
| Jiyuan      | -.56             | 16            | -.97535  | 17         | -.14784  | 10         |

Judging from the score of the first main factor, Zhengzhou, Luoyang and Kaifeng ranked the top two. This main factor represents the total social and economic scale. It can also be seen from Table 2 that these three cities have relatively high levels in the total retail sales of consumer goods, the number of industrial enterprises above designated size in each city, the proportion of fixed asset investment in transportation and storage, the number of trucks, and the indicators.

From the score of the second main factor, the top three are Nanyang, Zhengzhou and Xinyang. This factor represents the logistics infrastructure. Among them, Nanyang City has superior transportation and developed railways and highways, making it the forefront of Henan. Luoyang is only ranked fifteenth in the second main factor ranking, indicating that Luoyang should strengthen the construction of logistics infrastructure. Generally speaking, the higher the score, the stronger the logistics competitiveness of the region. A score greater than 0 indicates that the logistics competitiveness of the region is above the average level of the province, and a score less than 0 indicates that the logistics competitiveness of the region is lower than the average level of the province. From the data in Table 6, it can be seen that Zhengzhou has the strongest logistics competitiveness, and Zhumadian's logistics industry is the least competitive. Zhengzhou, Luoyang, Kaifeng, Pingdingshan and Anyang have high logistics competitiveness. The logistics competitiveness of the other 13 cities is lower than the provincial average.
3.3. Cluster analysis
This In order to deeply analyze the level and structure of the competitiveness of cities in Anhui Province, the two main factors extracted in Table 5 are used to further cluster analysis. The statistical software SPSS22.0 was used for systematic clustering, and the classification results of Table 7 were obtained.

| category | city |
|----------|------|
| 1        | Zhengzhou |
| 2        | Nanyang |
| 3        | Kaifeng, Luoyang, Pingdingshan, Anyang |
| 4        | Xinyang, Xinxiang, Hebi, Shangqiu, Jiyuan, Zhumadian |
| 5        | Puyang, Jiaozuo, Zhoukou, Xuchang, Sanmenxia, Luhe |

The first type of city: Zhengzhou. It can be seen from Figure 4 that Zhengzhou has the strongest performance in terms of logistics competitiveness, which is much higher than other cities, as a single category. As the city with the highest level of logistics development and the highest score for comprehensive development in Henan Province, Zhengzhou leads other cities in the province as the leader. As the provincial capital of Henan Province, Zhengzhou has a geographical location that links the east and the west, which facilitates the transfer of industries from the east and the export of resources from the west. As the most important transportation hub city in the central region, its superior location advantages and perfect logistics infrastructure have played an important supporting role in the development of Zhengzhou's modern logistics industry.

The second category of cities: Nanyang City. Nanyang City's economy is not developed, but it has relatively complete logistics infrastructure conditions. Kilometers, routes, cargo volume, and cargo turnover are all in the forefront of the province. It is located in a geographical location that connects east and west, and connects north to south. Convenient and good communication. This has created favorable conditions for the development of the logistics industry, and has great potential for future development.

The third category of cities: Kaifeng, Luoyang, Pingdingshan, Anyang. These cities are cities with strong comprehensive logistics competitiveness. On the one hand, they have superior geographical location, convenient transportation, and complete logistics infrastructure, which provide strong support for the development of modern logistics industry; on the other hand, the commercial and circulation industry is developed and is an important material in Henan Province. Distribution centers have many large-scale wholesale markets, which have a large demand for modern logistics and have the foundation for the development of modern logistics. The socio-economic development is relatively good, but the logistics demand and scale rank low, which ultimately weakens the overall logistics strength of the entire city.

The fourth category of cities: Xinyang City, Xinxiang City, Hebi City, Shangqiu City, Jiyuan City, Zhumadian City. The comprehensive capabilities of logistics competitiveness of these cities are average, and it is difficult for them to develop simultaneously in terms of social economy and logistics needs, and there are different differences. Among them, Xinyang City, Hebi City, and Xinxiang City rank high in terms of logistics demand and scale, but the local economic strength is not strong, and the advantages of logistics demand scale are not fully utilized. Similar problems exist in other cities, restricting the city's logistics development.

The fifth category of cities: Puyang City, Jiaozuo City, Zhoukou City, Xuchang City, Sanmenxia City, Luhe City. The social and economic strength of these cities is not outstanding, and some of them are poor in some aspects. Moreover, the logistics industry is still a new industry in these areas. The logistics market is not attractive and needs to be improved and further explored.

4. Conclusions and recommendations
Through the use of factor analysis and cluster analysis, the regional logistics competitiveness of 18 cities in Henan Province was comprehensively evaluated and ranked. The logistics competitiveness was
ranked from high to low, and the order was: Zhengzhou, Luoyang, Kaifeng, Pingdingshan, Anyang, Hebi, Xinxiang, Sanmenxia, Xuchang, Jiaozuo, Puyang, Luohe, Shangqiu, Nanyang, Zhoukou, Jiyuan, Xinyang, Zhumadian.

Based on the above analysis, the following suggestions are put forward: First, we must improve the uneven development of the logistics industry in various regions of Henan Province. Only 5 of the 18 cities in Henan Province have a positive overall score, and the remaining 13 cities have a negative score, which is lower than the average level of Henan Province. For areas with a low level of logistics development, the government should first cultivate leading enterprises in the logistics industry, play the role of leading enterprises, and promote the development of other logistics enterprises. Areas where conditions permit can also build logistics industry clusters to promote the coordinated development of enterprises in the clusters and achieve economies of scale. In short, each region should develop its own logistics industry in light of actual conditions and reduce the differences between regions. This is of great significance to the coordinated development of the economy of Henan Province.

Second, we must vigorously promote economic development. Generally speaking, the level of development of logistics competitiveness in each region corresponds to the level of regional economic development. Cities with better economic strength have stronger logistics competitiveness. The development of the logistics industry will in turn drive the further development of the economy. Therefore, all regions should speed up economic construction: in addition, increase investment in infrastructure. The 18 cities in Henan Province are all inland cities, and water transportation is underdeveloped. It can improve the transportation capacity of roads, railways, and air transportation, which is a modern logistics industry. Create conditions for the development of the country.

Finally, promote the optimization and upgrading of the logistics industry. Henan Province should introduce modern logistics technologies and concepts into the logistics industry, use GPRS systems, satellite systems, and the Internet in the logistics industry to improve the automation and intelligence level of the logistics industry, and increase circulation processing and logistics. Modern logistics services such as consulting, inventory control strategy recommendations, logistics scheme design and planning, promote the optimization and upgrading of the logistics industry as soon as possible.

References
[1] Mao Xiaoyi. Research on the Evaluation of Regional Logistics Industry Competitiveness [J]. Logistics Engineering and Management, 2021, 43(01): 52-53+61. (In Chinese)
[2] Tage Skjott-Larsen. Sten Wandel. Ulf Paulsson. Logistics in the Oresund Region after the Bridge [J]. European Journal of Operational Research. 2003 (1): 247-256.
[3] Dong Lina. Research on the Evaluation of Regional Logistics Competitiveness in the Yangtze River Delta [D]. Anhui University of Science and Technology, 2015. (In Chinese)
[4] Qu Jing. Research on the Evaluation of Regional Logistics Industry Competitiveness[D]. Harbin. Harbin University of Science and Technology, 2015. (In Chinese)
[5] Xia Fan. Comprehensive Evaluation and Comparison of Regional Logistics Competitiveness—Based on Factor Analysis and Cluster Analysis[J]. Chinese and Foreign Entrepreneurs, 2018, (12): 47-50. (In Chinese)
[6] Henan Provincial Bureau of Statistics. Henan Provincial Statistical Yearbook (2019) [M]. Beijing: China Statistics Press, 2020. (In Chinese)
[7] Zhong Xiaoyan, Ou Weiqiang, Chen Shengzhe. An Empirical Analysis of the Competitiveness Evaluation of Cold Chain Logistics Enterprises[J]. Journal of Liaoning University of Technology (Social Science Edition), 2021, 23(01): 48-52. (In Chinese)