BULGARIAN ACADEMY OF SCIENCES

CYBERNETICS AND INFORMATION TECHNOLOGIES • Volume **14**, No 4

Sofia • 2014 Print ISSN: 1311-9702; Online ISSN: 1314-4081
DOI: 10.1515/cait-2014-0008

Evaluation of the Software Industry Competitiveness in Jilin Province Based on Factor Analysis

Pinchao Meng[1], *Weishi Yin*[1], *Yanzhong Li*[2]

[1]*College of Science, Changchun University of Science and Technology, Changchun, 130022 China*
[2]*Beihua University, Jilin, 132012 China*
Emails: mengpc@cust.edu.cn yinweishi@foxmail.com lyz@cust.edu.cn

Abstract: *In this paper 12 economic indices of the software industry in 30 cities/provinces in China are used to set up an evaluation system for the competitiveness of the regional software industry. By using the statistical analysis method of factor analysis, an evaluation model of the comprehensive competitiveness of the software industry for each city/province is built. Taking Beijing and Shanghai as examples, the comprehensive competitiveness and problems of the software industry in Jilin province are compared and analyzed.*

Keywords: *Software industry, competitiveness, factor analysis.*

1. Introduction

As a strategic leading industry in 21-st century, the software industry is crucial to the electronic information industry and plays a decisive role in the information industry development, thus becoming one of the measurements to assess the Comprehensive National Power. N o w a k and G r a n t h a m [1] characterized the major barriers to success for small market software entrepreneuiral ventures based on the software industry in California, United States, and further built and verified their econometric model. D e h u a J u [2] studied the software industry in Ireland and demonstrated that the reason for Ireland holding a dominant position in the software market lays in its high quality software and the advantage of human resources management. D a y a s i n d h u [3] has developed a dynamic theoretical framework for global competitiveness and implied that the organizations in the Indian software industry have created trust and encouraged the inter-organizational relationships. Deependra [4] studied the development of the Indian software industry comprehensively. Chen [5] proposed options to enhance the software industry development healthily with respect to the major existing problems of the software industry in China. Liu [6] compared the software industry in China to that in India based on Porter's "The Competitive Advantage of Nations".

The development of the software industry in Jilin Province plays a pivotal role in developing old industrial bases. In our research the competitiveness of the software industry in Jilin province, Shanghai and Beijing was compared. Moreover, the factor analysis method was used to evaluate the comprehensive competitiveness of the software industry in Jilin Province and further the overall situation and current problems were analyzed.

2. The evaluation indices of the competitiveness in the software industry

Starting from the formation of the software industry competitiveness, we have built our evaluation system based on Porter's diamond model, IMD world competitiveness index and the evaluation system of the Chinese Enterprise Confederation. Moreover, we have also considered a recent evaluation system study on the software industry competitiveness in China and combined the features of the software industry and data acquisition. Our evaluation system for the competitiveness of the regional software industry includes the three first level indices (industry input, industry output and market performance) and 12 second level indices.

Table 1. Evaluation system for the competitiveness of the software industry in Jilin Province

General objective	First level indices	Second level indices
The evaluation system for the competitiveness of the software industry in Jilin Province	Industry input	Enterprise number x_1
		Number of employees x_2
		R&D staff x_3
		Bachelor degree and above proportion x_4
		R & D funds x_5
		R&D input intensity x_6
	Industry output	Business income x_7
		Industrial added value x_8
		Total profit x_9
	Market performance	Market share x_{10}
		Export ratio x_{11}
		Product exports x_{12}

3. Factor analysis

Factor analysis is proposed by the British psychologist C. Speannan and has been widely applied in the field of economics, management science and sociology. Based on the analysis of a few factors, factor analysis is a method to study the connection and quantity relationship between the factors and the original variables and to detect the internal structure of the factors. Further factor analysis can be used to construct a comprehensive evaluation function and perform evaluation based on the scores obtained via such function. The steps are the following:

3.1. Hypothesis test

Generate a correlation coefficient matrix and perform a test of significance to determine whether it is feasible to perform factor analysis for the original variables.

3.2. Normalization

Let the sample data matrix be $Z = (Z_{ij})_{m\times n} = \{z_1, z_2, \cdots, z_n\}$, where there are n variables and m observations. Let $X = (X_{ij})_{m\times n} = \{x_1, x_2, \cdots, x_n\}$ be the matrix after normalization, and the mean and standard deviation for each observation be 0 and 1, respectively.

The transformation equation for normalization is

$$X_{ij} = \frac{Z_{ij} - \overline{z_j}}{S_j}, \quad i = 1, 2, 3, \cdots, n, \qquad j = 1, 2, 3, \cdots, p,$$

$$\overline{z_j} = \frac{1}{n}\sum_{i=1}^{n} Z_{ij}, \quad S_j^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Z_{ij} - \overline{z_j})^2.$$

3.3. Generating a sample correlation matrix, eigenvalues and eigenvectors

Let the sample correlation matrix be $R=(r_{ij})_{m\times n}$, and $r_{ij}=r_{ji}$ and $r_{ii}=1$, then R is a symmetric matrix with ones in the main diagonal. Let the eigenvalues and eigenvectors for R be $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p$ and $u_1, u_2, \ldots, u_p$, respectively, and $u_i = (u_{i1}, u_{i2}, \cdots, u_{ip}), \quad i = 1, 2, \cdots, p$.

3.4. Extracting principal components

Based on the eigenvalues and eigenvectors, we compute the Variance Contribution Rate (VCR) $w_i = \lambda_i / \sum \lambda_j$ and the factor loading of X_i on F_i, $a_{ij} = u_{ij}\sqrt{\lambda_j}$, and extract k principal components according to the criteria that the accumulation of VCR ≥ 85%. Then we perform orthogonal rotation of the loading matrix, making the matrix as close as possible to the direction of +1, –1 or 0, and write down the factor equations:

$$\begin{cases} X_1 = a_{11}F_1 + a_{12}F_2 + \cdots + a_{1k}F_k, \\ \qquad \cdots, \\ X_n = a_{n1}F_1 + a_{n2}F_2 + \cdots + a_{nk}F_k. \end{cases}$$

The eigenvalue λ_i for the i-th component F_i measures the variance of this component. The bigger the variance, the more contribution it has to the overall variance. The contribution rate w_i is the percentage of the explanatory importance of the corresponding factor with respect to all variables. The k components of the corresponding eigenvector u_i with respect to the eigenvalue λ_i are the coefficients for the k normalized variables of the corresponding F_i, and the absolute values and signs indicate the correlation strength and direction between the principal component and the corresponding variable.

3.5. The interpretation of the principal components

The coefficients a_{ij} of the factor loading matrix indicate that the relative high loadings are distributed among the components regularly. We interpret the components according to their actual meanings.

3.6. Calculating the factor scores from the loading matrix A and building the factor evaluation model

Calculate the factor scores from the obtained loading matrix A:

$$F_i = \left(\sum_j a_{ij} x_{ij} \right) / \sum_j a_{ij}.$$

Weighted with w_i, the VCR of each principal component and based on the above factor scores, the comprehensive evaluation model is built:

$$F = \sum_{i=1}^{k} w_i F_i.$$

F is the comprehensive score for the competitiveness of the regional software industry; w_i is the weight for the i-th component (VCR for the i-th component); F_i is the factor score for the i-th component.

3.6.1. Establishing of the evaluation model of the competitiveness of the regional software industry

a) Kaiser-Meyer-Olkin (KMO) test and Bartlett's test of sphericity

We performed KMO test and Bartlett's test of sphericity after normalizing the 12 economic indices for 30 cities/provinces in 2012 in China, and the results are shown in Table 2.

Table 2. Kaiser-Meyer-Olkin and Bartlett's test

KMO measure of sampling adequacy		0.848
Bartlett's test of sphericity	Approx. Chi-Square	862.128
	df	66
	Sig.	0.000

As it can be seen from the table, the first row shows that the KMO test value is 0.848, and the second row shows the Bartlett's test result rejecting the null hypothesis, thus the factor analysis is significant for the 12 indices and it is feasible to perform the factor analysis.

Table 3. Communalities

Variable	Initial	Extraction	Variable	Initial	Extraction
x_1	1.000	0.959	x_7	1.000	0.954
x_2	1.000	0.978	x_8	1.000	0.937
x_3	1.000	0.895	x_9	1.000	0.912
x_4	1.000	0.816	x_{10}	1.000	0.955
x_5	1.000	0.948	x_{11}	1.000	0.840
x_6	1.000	0.458	x_{12}	1.000	0.851

Table 3 demonstrates the initial communality and the re-generated communality after extracting 2 factors for the 12 indices used in our analysis. The common factor analysis exhibits that the majority of the indices shares more than 90% communality. The above evaluation demonstrates the tight internal structural relationship between the extracted principal components and the indices and thus it is suitable for factor analysis.

b) Extracting the principal components

Calculate the eigenvalues, VCRs and accumulation of VCR for the principal components before and after rotation (Table 4).

Table 4. Total variance explained

No	Initial eigenvalues			Extraction sums of squared loadings			Rotation sums of squared loadings		
	Total	Variance, %	Cumulative, %	Total	Variance, %	Cumulative, %	Total	Variance, %	Cumulative, %
1	9.337	77.812	77.812	9.337	77.812	77.812	9.175	76.459	76.459
2	1.167	9.722	87.533	1.167	9.722	87.533	1.329	11.074	87.533
3	0.770	6.420	93.953						
4	0.419	3.490	97.443						
5	0.121	1.011	98.455						
6	0.083	0.695	99.150						
7	0.059	0.489	99.639						
8	0.025	0.209	99.848						
9	0.010	0.084	99.931						
10	0.006	0.052	99.983						
11	0.002	0.016	100.000						
12	5.037×10^{-5}	0.000	100.000						

In Table 4, the first column is a factor index, and starting from the second column, every three columns are grouped together. In each group, the columns are eigenvalues, VCRs and accumulations of VCR in sequence. The three groups of data demonstrate the initial solution, the solution and the final solution for factors after rotation. We can see that the eigenvalues for the 2 principal components after rotation are 9.175 and 1.329 respectively, the VCRs are 76.459% and 11.074% respectively, and the accumulation of VCR is 87.533%. The two components can explain as much as 87.533% variance of the 12 indices, which means that the two components contain major information of the original variables and are eligible to evaluate the competitiveness of the regional software industry, as main factors.

c) The analysis of the factor loading matrix and interpretation of the principal components

Table 5. The factor loading matrix after rotation

Variable	Component		Variable	Component	
	1	2		1	2
x_1	0.962	0.183	x_7	0.976	0.048
x_2	0.986	0.083	x_8	0.968	0.029
x_3	0.917	0.234	x_9	0.953	0.051
x_4	–0.125	0.895	x_{10}	0.976	0.044
x_5	0.971	0.071	x_{11}	0.880	0.258
x_6	0.346	0.581	x_{12}	0.914	0.125

Table 5 demonstrates the result of rotating the factor loading matrix with the maximum variance method. Using the principal components analysis to perform a rotation with maximum variance method, we get a relatively high loading on the first factor of X_1, X_2, X_3, X_5, X_7, X_8, X_9, X_{10}, X_{11}, X_{12}, where ten variables could reflect the general competitiveness of industry, and define the first principal component as a general strength factor. We get a relatively high loading on the second factor of X_4, X_6 and define the second factor as an innovation strength factor. Thus we have the factor equation:

$$\begin{cases} X_1 = 0.962F_1 + 0.183F_2 \\ X_2 = 0.986F_1 + 0.083F_2 \\ X_3 = 0.917F_1 + 0.234F_2 \\ X_4 = -0.125F_1 + 0.895F \end{cases}, \quad \begin{cases} X_5 = 0.971F_1 + 0.071F_2 \\ X_6 = 0.346F_1 + 0.581F_2 \\ X_7 = 0.976F_1 + 0.048F_2 \\ X_8 = 0.968F_1 + 0.029F_2 \end{cases}, \quad \begin{cases} X_9 = 0.953F_1 + 0.051F_2 \\ X_{10} = 0.976F_1 + 0.044F_2 \\ X_{11} = 0.880F_1 + 0.258F_2 \\ X_{12} = 0.914F_1 + 0.125F_2 \end{cases}.$$

d) Calculating the factor scores and set up the evaluation model

The factor score function can be deduced from Table 5:

$$\begin{cases} F_1 = (0.962X_1 + 0.986X_2 + 0.917X_3 + 0.971X_5 + 0.976X_7 + 0.968X_8 + 0.953X_9 + \\ \quad + 0.976X_{10} + 0.880X_{11} + 0.914X_{12})/(0.962 + 0.986 + 0.917 + 0.971 + 0.976 + \\ \quad + 0.968 + 0.953 + 0.976 + 0.880 + 0.914), \\ F_2 = (0.895X_4 + 0.581X_6)/(0.895 + 0.581). \end{cases}$$

According to the weight for each principal component and the corresponding score, we have the evaluation model of the competitiveness of the regional software industry:

$$F = w_1F_1 + w_2F_2 = 0.873F_1 + 0.127F_2,$$

where $w_i = \lambda_i / \sum \lambda_j$, λ_i is the variance for *i*-th component, F is the comprehensive score for the competitiveness of the regional software industry, and F_1, F_2 are the factor scores for the first and second principal components, respectively.

3.6.2. Evaluation of the comprehensive competitiveness of the software industry in Jilin

Given our newly-built evaluation model, we have chosen two typical regions of China, Beijing and Shanghai as comparison targets and have compared them to Jilin. Further on we have evaluated the comprehensive competitiveness of the software industry and studied the competitiveness strength in Jilin Province.

F_1, F_2 and F values were calculated after normalizing the raw data for 30 different cities/provinces and are shown in Fig. 6.

Table 6. F_1, F_2 and comprehensive competitiveness scores (F) for Beijing, Shanghai and Jilin

Region	F_1	F_2	F
Beijing	1.49	0.18	1.32
Shanghai	0.98	0.60	0.93
Jilin	–0.44	0.23	– 0.36

Our results have only demonstrated the relative differences rather than the absolute values for the competitiveness of the software industry for each city/province.

Altogether, the comprehensive competitiveness scores for the software industry in Jilin Province, Beijing and Shanghai are – 0.31, 1.16 and 0.81 respectively, exhibiting a big gap between Jilin Province and Beijing/Shanghai. With respect to the two components, the VCRs for the first factor and the second factor are 0.873 and 0.127 respectively, indicating the leading role of the first factor on the competitiveness of the regional software industry. As shown in Fig. 6, the first factor score is –0.44 of Jilin Province, which is much smaller than that of Beijing (1.49) and Shanghai (0.98). This is due to the professional personnel shortages and insufficient funding, as well as the relative weak competitiveness in marketing performance and technological innovations. The score for the second factor is 0.23 of Jilin Province, which is higher than Beijing (0.18) and lower than Shanghai (0.60). This is the result of the advantages of IT personnel and the emphasis on team building and personnel training in Jilin Province, however still accompanied with insufficient funding. As a result, the relative weak comprehensive competitiveness of the software industry in Jilin Province is due to the relative low score for the first factor, which acts as a leading factor.

4. Conclusions

A case study was performed for the competitiveness of the software industry in Jilin Province, using factor analysis. By comparing and analyzing the related data for the software industry in 30 cities/provinces in China, the relative development level of the software industry in Jilin Province was obtained. After further comprehensive evaluation, a gap was recognized between the software industry in Jilin and in the developed cities/provinces in China, and light is shed on the government strategy for industry development.

References

1. N o w a k, M. J., C. E. G r a n t h a m. The Virtual InCubator: Managing Human Capital in the Software Industry. – Research Poliey, Vol. **29**, 2000, No 10, 125-134.
2. J u, D. China's Budding Software Industry. – IEEE Software, 2001, 5-6.
3. D a y a s i n d h u, N. Embeddedness. Industry Clusters and Global Competitivness: A Case Study of the Indian Software Industry. – Technovation, Vol. **22**, 2002, 551-560.
4. D e e p e n d r a, M. India's Software Industry. – IEEE Software, 2001, 1-2.
5. C h e n, C. WTO-ITA and China IT Industry Development. Beijing, Beijing University of Posts and Telecommunications Press, 2001, 162-189.
6. L i u, L. The Comparative Research on the Development of the Software Industry in China and India. Changsha, Hunan University, 2004.
7. Z o o k, C. , J. A l l e n. Profit from the Core: Growth Strategy in an Era of Turbulenee. HBS Press Book, February 2001.