

# General Evaluation of Economic Development Level Based on Principal Component Analysis

Zhang Hongmin

School of Finance, Harbin University of Commerce

## Abstract

In assessing a country's degree of economic development, the various aspects of social production should be taken into account, and we look at the combined effect of productivity at last. In order to an objective and scientific analysis of China's economic development, this paper applies principal component analysis on China's 31 provinces, municipalities and autonomous regions (not including Hong Kong, Macao and Taiwan) to analysis the economic development level and work out a comprehensive analysis and evaluation by using the SPSS software to make computer computation. Accordingly, the results highlighted the characteristics and advantages of major provinces in the economic development process, and provide a theoretical basis to balanced development of China.

**Keywords:** Principal Components Analysis, Economic development, Comprehensive Evaluation

## 1. Introduction

To describe and evaluate the economic development of society, the most ideal method is to find out an all-inclusive criteria system for evaluation, whose testing result can reflect most or the overall information of social and economic development. Since 1960s, some international organization, some national or regional functional department and some researchers have ever proposed various criteria system for evaluation. For our country, the systematical study on criteria system for evaluating the social development got a late start,

but gained a rapid development. Since 1980s, some government department, research facilities and researchers have designed some evaluation systems of social indicators, for example, Tang Xiaodong used a functional model consisting of 21 criteria variables to evaluate the social and economic development of our country, but there was a largest disadvantage for this model, it didn't take account of all the factors which can reflect economic development, and therefore the expecting result can't be obtained. So far a more perfect and objective evaluation system of social indicators hasn't been formed. To represent an overall and objective description of the social development of all regions in our country, this paper, based on the research results at home and abroad and made a further amendment and enrichment, brought in consumer price index and commodity retail price index for evaluation system of social indicators for the first time and proposed a new evaluation system of social indicators. In real economic problems, different economic variables are relevant for a certain degree, for example, workers' average wage and consumption level definitely are relevant, which added the complication of analyzing the problem, thus a simplified method is needed. The principal component analysis method can use less indicators to replace the method which used more indicators, moreover the less indicators can reflect what the original indicators showed, thus radically avoid information overlapping between indicators and greatly simplified indicator structure of original indicator system.

## 2. Indicator identification

Based on the principle that the selected indicators should reflect the region's economic

development, and data acquisition situation, in addition, according to the problem studied, this paper selected 8 main indicators which can reflect economic development: Gross Domestic Product( $X_1$ ), household consumption level ( $X_2$ ), investment in infrastructure( $X_3$ ), workers' average wage( $X_4$ ), Consumer Price Index( $X_5$ ), retail price index( $X_6$ ), cargo turnover ( $X_7$ ), gross industrial output value( $X_8$ ). The indicator data were got from China Statistical Yearbook of 2011 and other relevant organized data and they were normalized and then the software of SPSS was made use of to analyze the main component and the variance explained table showed as below.

Tab.1 variance explained table

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.742	46.781	46.781	3.742	46.781	46.781
2	2.392	29.901	76.682	2.392	29.901	76.682
3	.740	9.246	85.928	.740	9.246	85.928
4	.480	6.004	91.933			
5	.438	5.472	97.405			
6	.142	1.770	99.174			
7	.059	.743	99.918			
8	.007	.082	100.000			

From table 1, it can be seen that the contribution of 3 characteristic values has reached 85.928% in total. This showed that the information reflected by former 3 factors accounted for more than 85.928% of all the information. Therefore, the former 3 factors can serve as the main component and the analysis result also showed this point.

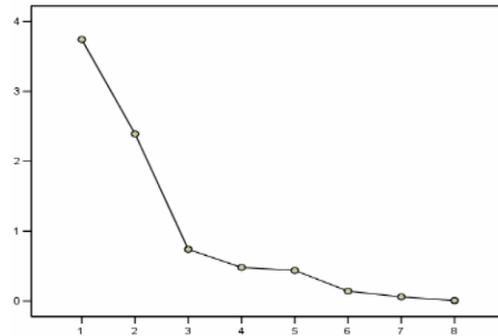


Fig 1 Scree Plot

Table 1 showed the analysis result—Scree Plot, the significant turning point is 3, so the former 3 factors can include most of the information. The contribution of the former 3 factors accounts for 85.918% of the total variance and this shows that selecting the former 3 common factors is more appropriate.

Tab.2 Component Matrix

	component		
	1	2	3
Gross Domestic Product	.815	.555	-.119
Household Consumption Level	.766	-.493	.196
Investment in Infrastructure	.785	.559	-.140
Workers' Average Wage	.604	-.572	.016
Consumer Price Index	-.315	.597	.666
Retail Price Index	-.397	.721	-.004
Cargo Turnover	.761	-.182	.458
Gross Industrial Output Value	.822	.541	-.115

From table 2, it can be seen that Gross regional domestic product and gross industrial output value has a larger load on the first main component, that is to say, its correlation index with the first principal component is bigger; the investment in infrastructure has a larger load on the second principal component, that is to say, it has more correlation with the second principal component; consumption price index has a larger load on the third component, that is to say, it has a great relation with the third principal component. Thus the principal component was named as: The first principal component is output principal component; the second principal component is construction investment principal component; the third principal component is consumption price principal component.

Tab.3 Component Score Coefficient Matrix

component <sup>o</sup>	component <sup>o</sup>		
	1 <sup>o</sup>	2 <sup>o</sup>	3 <sup>o</sup>
Gross Domestic Product <sup>o</sup>	.218 <sup>o</sup>	.232 <sup>o</sup>	-.161 <sup>o</sup>
Household Consumption Level <sup>o</sup>	.205 <sup>o</sup>	-.206 <sup>o</sup>	.265 <sup>o</sup>
Investment in Infrastructure <sup>o</sup>	.210 <sup>o</sup>	.234 <sup>o</sup>	-.189 <sup>o</sup>
Workers' Average Wage <sup>o</sup>	.161 <sup>o</sup>	-.239 <sup>o</sup>	.022 <sup>o</sup>
Consumer Price Index <sup>o</sup>	-.084 <sup>o</sup>	.250 <sup>o</sup>	.901 <sup>o</sup>
Retail Price Index <sup>o</sup>	-.106 <sup>o</sup>	.301 <sup>o</sup>	-.005 <sup>o</sup>
Cargo Turnover <sup>o</sup>	.203 <sup>o</sup>	-.076 <sup>o</sup>	.620 <sup>o</sup>
Gross Industrial Output Value <sup>o</sup>	.220 <sup>o</sup>	.226 <sup>o</sup>	-.155 <sup>o</sup>

According to this table and the observed value of variances, factors score can be calculated, the value of principal Component Score Coefficient Matrix is got from principal component divided by corresponding characteristic root. So far, 8 evaluation indicators were transformed into 3 comprehensive evaluation indicators which had typical economic connotation by principal component meaning.

**3. Comprehensive analysis and evaluation of economic development**

In order to investigate every city and evaluate them, 3 principal component were expressed as a linear combination of the 8 indicators which is score function

$$Z_1 = 0.218 * \sqrt{3.742} \frac{X_1 - 437223}{3588.6008} + L + 0.22 * \sqrt{3.742} \frac{X_8 - 1831.535}{1770.4038}$$

$$Z_2 = 0.232 * \sqrt{2.392} \frac{X_1 - 437223}{3588.6008} + L + 0.226 * \sqrt{2.392} \frac{X_8 - 1831.535}{1770.4038}$$

$$Z_3 = -0.161 * \sqrt{0.74} \frac{X_1 - 437223}{3588.6008} + L - 0.155 * \sqrt{0.74} \frac{X_8 - 1831.535}{1770.4038}$$

The above function can be regarded as evaluation model of urban economic developments, the scores can be got by inserting the standardization data of 31 provinces and cities as the table 4 showed.

To illustrate one point, the principal component score for some city is negative number because of the standardization process of original data in calculation---regarding the average level of every economic indicator as zero. Consequently, the negative number only

showed that its economy is less developing than the average development of the 31 provinces and cities. Putting the score of Z1、Z2、Z3 in three dimensional system of coordinate as table 2 shows. From load matrix and the expression Z1、Z2、Z3, the larger the value of Z1、Z2、Z3 is, the stronger comprehensive capacity of the province or city it represents.

Tab.4 scores Table

	Beijing <sup>o</sup>	Tianjin <sup>o</sup>	Hebei <sup>o</sup>	Shanxi <sup>o</sup>	Neimeng- Gu <sup>o</sup>	Liao- Ning <sup>o</sup>	Jilin <sup>o</sup>	Heilong- Jiang <sup>o</sup>	Shanghai <sup>o</sup>	Jiangsu <sup>o</sup>
Z <sub>1</sub> <sup>o</sup>	2.629 <sup>o</sup>	1.949 <sup>o</sup>	0.494 <sup>o</sup>	-1.277 <sup>o</sup>	-1.011 <sup>o</sup>	0.710 <sup>o</sup>	-1.18 <sup>o</sup>	-0.124 <sup>o</sup>	5.064 <sup>o</sup>	3.352 <sup>o</sup>
Z <sub>2</sub> <sup>o</sup>	-3.416 <sup>o</sup>	-3.133 <sup>o</sup>	1.576 <sup>o</sup>	0.489 <sup>o</sup>	0.273 <sup>o</sup>	-0.103 <sup>o</sup>	-0.024 <sup>o</sup>	-0.127 <sup>o</sup>	-3.410 <sup>o</sup>	2.045 <sup>o</sup>
Z <sub>3</sub> <sup>o</sup>	1.907 <sup>o</sup>	1.614 <sup>o</sup>	1.251 <sup>o</sup>	0.610 <sup>o</sup>	1.094 <sup>o</sup>	0.709 <sup>o</sup>	-0.150 <sup>o</sup>	-0.592 <sup>o</sup>	1.367 <sup>o</sup>	-1.212 <sup>o</sup>
Zhejiang <sup>o</sup>	Anhui <sup>o</sup>	Fujian <sup>o</sup>	Jiangxi <sup>o</sup>	Shan- Dong <sup>o</sup>	Xinjiang <sup>o</sup>	Henan <sup>o</sup>	Hubei <sup>o</sup>	Hunan <sup>o</sup>	Guang- Dong <sup>o</sup>	Guang- Xi <sup>o</sup>
2.965 <sup>o</sup>	-1.137 <sup>o</sup>	0.362 <sup>o</sup>	-1.089 <sup>o</sup>	3.224 <sup>o</sup>	-1.899 <sup>o</sup>	0.291 <sup>o</sup>	-0.653 <sup>o</sup>	-0.949 <sup>o</sup>	4.275 <sup>o</sup>	-1.212 <sup>o</sup>
1.085 <sup>o</sup>	1.150 <sup>o</sup>	-0.895 <sup>o</sup>	-0.159 <sup>o</sup>	2.387 <sup>o</sup>	-1.223 <sup>o</sup>	2.004 <sup>o</sup>	1.829 <sup>o</sup>	1.142 <sup>o</sup>	1.581 <sup>o</sup>	-0.114 <sup>o</sup>
0.191 <sup>o</sup>	0.440 <sup>o</sup>	-0.534 <sup>o</sup>	-0.758 <sup>o</sup>	-0.516 <sup>o</sup>	-1.168 <sup>o</sup>	0.587 <sup>o</sup>	0.852 <sup>o</sup>	1.261 <sup>o</sup>	2.265 <sup>o</sup>	-0.338 <sup>o</sup>
Hainan <sup>o</sup>	Chong- Qing <sup>o</sup>	Sichuan <sup>o</sup>	Guizhou <sup>o</sup>	Yunnan <sup>o</sup>	Xizang <sup>o</sup>	Shanxi <sup>o</sup>	Gansu <sup>o</sup>	Qinghai <sup>o</sup>	Ningxia <sup>o</sup>	
-1.799 <sup>o</sup>	-0.866 <sup>o</sup>	-0.328 <sup>o</sup>	-2.132 <sup>o</sup>	-1.163 <sup>o</sup>	-2.944 <sup>o</sup>	-1.346 <sup>o</sup>	-1.651 <sup>o</sup>	-2.177 <sup>o</sup>	-1.887 <sup>o</sup>	
-1.062 <sup>o</sup>	-0.878 <sup>o</sup>	0.963 <sup>o</sup>	-0.278 <sup>o</sup>	-0.248 <sup>o</sup>	-2.442 <sup>o</sup>	0.499 <sup>o</sup>	-0.434 <sup>o</sup>	-0.074 <sup>o</sup>	-0.837 <sup>o</sup>	
1.436 <sup>o</sup>	-1.037 <sup>o</sup>	0.036 <sup>o</sup>	-0.268 <sup>o</sup>	-0.296 <sup>o</sup>	-0.462 <sup>o</sup>	0.269 <sup>o</sup>	-0.312 <sup>o</sup>	0.811 <sup>o</sup>	0.428 <sup>o</sup>	

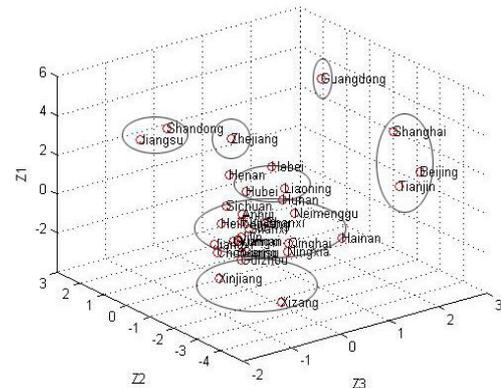


Fig2 Score spread map

According to table 2, it can be seen that above the middle is Guangdong province and its gross product, infrastructure investment and consumption level is much higher relatively; while below the middle is Tibet and its consumption level is on par with average consumption level of all the country; cities on the upper left of the three dimensional coordinate have higher gross product and infrastructure investment such as Jiangsu and Shandong; Cities on the upper right have a larger gross product and a higher consumption level such as Shanghai, Beijing and Tianjin. Based on the three aspects—gross domestic product and infrastructure investment and consumption level, evaluations were made on the regional economic development of our country: 31 provinces and

cities can be divided into 7 kinds: Guangdong is one kind. It stands on the right above of Table 2, which illustrates that the gross domestic value, infrastructure investment and consumption level is at a higher level. It is the best place for investors. Jiangsu and Shandong are one kind. They are on the upper left of Table 2. Their gross domestic product and infrastructure investment is much higher but consumption level is relatively not quite high; Hebei, Henan and Liaoning are one kind. Their gross domestic product is at the same level with the average gross domestic product of all the country but their infrastructure investment is much higher, which proves that infrastructure development is improving. Zhejiang is one kind. Its gross domestic product is higher than Hebei, Henan and Liaoning, but its infrastructure investment and consumption level is almost at the same level with the three provinces; Shanghai, Beijing and Tianjin are one kind. Their gross domestic product and consumption level is relatively higher but their infrastructure investment is relatively lower, which illustrates that the infrastructure development is more perfect; Hainan, Guizhou, Qinghai, Guansu, Xinjiang, Ningxia and Tibet are one kind. Their consumption level is almost the same with the average consumption level of all the country but their gross domestic product is far behind the average gross domestic level of our country and their infrastructure investment is much higher than the average level of our country's infrastructure development, which proves that government gave high priority to the economic development of these provinces and cities. Other provinces are one kind.

#### 4. Conclusion

This paper introduced in principal component analysis to the comprehensive

evaluation and analysis of urban economic development, took advantage of function of principal component and proposed a more precise method to evaluate urban economic development. This has a realistic significance for investors, policy formulation department of national economy and the city itself. Because of the standardization process for indicators in this method, it avoided the influences on analysis result from different dimensions and meanwhile the principal component method successfully eliminated the information overlapping between variances and avoids the deficiency of identifying indicators subjectively, which made the analysis result objective and clear. In addition, the convenient software can be applied in the whole process, such as SPSS and SAS.

#### References

- [1] Noorbakhsh, Farhad. Human Development and Regional Disparities in china: A Policy Model [J]. Journal of International Development, 2007, Vol. 14: 927 - 949.
- [2]Qiaofeng, Yaojian, the application of Time-series of generalization Principal Components Analysis on the dynamic description of economic development[J] Application of Statistics and Management
- [3] GAVIN W J, VISARIA P. Urbanization in large developing countries. China, Indohesia, Brazilamd India[M].Oxford: Clarendon Press, 1997.
- [4] LIAN Yuming. China cities report [R]. Beijing: China times and economics Press, 2004.
- [5]Tang Xiaodong, the evaluation system and empirical study on indicators of China's urban development
- [6]State statistic bureau, China's statistic yearbook[M]. Beijing: China statistic press,2011