Assessment of Regional Higher Education Development in China Based on Factor Analysis

WU JIAYU 伍佳昱 11920132204811 Instructed by Professor LIU YANAN

Multivariate Statistical Analysis, Double Degree in Mathematical Statistics

The Wang Yanan Institute for Studies in Economics, Xiamen University, China

Abstract: Since the dawn of the 21th century, the scale and quality of China's higher education has experienced a dramatic expansion in the light of industry upgrading as well as technology advancement. However, it is notable that the development of higher education in China has been characterized by regional unbalance. A factor analysis is hence conducted based on 2014 data concerning higher education situations in China's 31 provinces, autonomous regions or municipalities, in order to make a reasonable assessment of current higher education status in different regions, with a view to countermeasures for future school-running practice.

Key words: higher education higher education scale higher education quality

Ⅰ. Introduction

Higher education is an optional final stage of formal learning that occurs after secondary education. It not only prepares talents and elites for further career development before awarding them academic degrees or professional certifications, but also injects momentum into the boost of economy and culture for the whole society.

China has experienced a dramatic growth in terms of educational level. According to 2014 China Statistical Yearbook， the illiteracy rate has dropped from 80% in 1949 to 4.92% in 2014, while the transition from elite to mass higher education has been accomplished in 2002 as the enrollment ratio exceeded 15% (Trow & Burrage, 2010). Meanwhile, the spatial pattern of higher education level in China has long been highly unbalanced despite the surge in overall scale and standard. The education development in eastern China is generally more advanced than in the west, while the strip along Beijing-Guangzhou forms an “axis” where the education level is prominently higher (Wang, 2012).

This paper endeavors to construct a measure for the higher education status for districts based on 2014 data with a factor analysis approach, and it further analyzes the current situations in China’s 31 districts with a view to countermeasures for future school-running practice.

In the first part, the background and the significance of the research topic are discussed, and then an introduction to the organization of this paper is given. In the second part, variables and observations selected will be explained accompanying a descriptive overview. Then a factor analysis is performed in Part Three to achieve an assessment measure of current regional higher education development. At last, analyses and countermeasures are given in the fourth conclusion part.

Ⅱ. Variables Selection and Data Description

The focus of this paper is higher education in China, i.e., the optional professional

learning that occurs after secondary education delivered at universities, academies and other college-level institutions granting academic or professional degrees. Based on considerations of data availability and statistical caliber consistency, following five variables are selected as of interest, before the latest data in China's 31 provinces, autonomous regions or municipalities are retrieved from 2014 China Statistical Yearbook.

1. Number of Students per 100000 population - nsp

   The coverage of higher education is directly reflected by the number of students undergoing higher education relative to regional population. The higher this figure, the larger the chance a secondary school graduate enter into higher education.

2. Number of Educational Personnel per 100000 population - ntp

   Educational personnel, including educational personnel, administrative personnel and and other supporting staff, is crucial to higher education organization as it defines the amount of higher education resources available in the district.

3. Number of Institutes per 100000 population - nip

   Higher education institutes including universities, colleges and other college-level institutes serve as the venue for academic researches as well as for higher education. Moreover, higher education enrollment often give priority to local students, which makes a larger number of local institutes even more of a regional advantage.

4. Student-Teacher Ratio (Number of Teacher = 1) - str

   Student-Teacher Ratio = the number of students / the number of teachers, which can be approximately interpreted as the number of students each teacher is responsible for on average. As higher education is a form of elite education with focus on specialization and individualization, the lower the S-T ratio, the higher the quality of education is likely to be.

5. Senior Teachers Ratio (%) - snt

   Teacher with senior titles (senior and sub-seniors included), are precious to

higher education due to their specialization in their own fields and their experiences in teaching practice. The proportion of senior teachers among all full-time educational staff effectively measures the standard of higher education.

Research Object	Variables		No. of Observations	Mean	Median	SD
Higher Education	No. of Students per 100000	nsp	31	2552.307734	2355.959226	822.3738949
	No. of Personnel per 100000	ntp	31	224799.0041	187406.1159	125249.014
	No. of Institutes per 100000	ntp	31	0.02498121	0.024187307	0.007613614
	Student-Teacher Ratio	str	31	17.52438429	17.76517048	1.161900834
	Senior Teachers Ratio (%)	snt	31	41.70591626	40.3318647	5.922218582

Table 1. Descriptive statistics of variables

Ⅲ. Factor Analysis

1. Firstly, try a principal component factor analysis. Conduct a spectral factorization and choose number of factor = 2 or = 3 based on eigenvalues and proportions of sample variance explained as given in Table 2.

Factor	Eigenvalue	Cumulative Proportion of total Sample Variance explained
1	3.46053281	0.692106562
2	1.09757199	0.91162096
3	0.22453704	0.956528368
4	0.16759544	0.990047456
5	0.04976273	1.000000002

Table 2. Eigenvalues and proportions of sample variance explained

2. Then perform analysis with 2 and 3 factors and conclude that both are sufficient according to Bartlett test. Try a vraimax rotation, and find that 2 factor is more suitable for factor interpretation. Perform a maximum likelihood factor analysis including a vraimax rotation, and compare the solution with that from principal component in Table 3. It can be observed that the loadings group in the same manner.

Variable	Principle component				Maximum likelihood			
	F1	F2	Communalities	Specific variance	F1	F2	Communalities	Specific variance
nsp	-0.9707169	-0.0298781	0.943184	0.05681599	0.9746174	0.05707665	0.9531368	0.04686318
ntp	-0.94866076	-0.2282696	0.9520643	0.04793573	0.93480772	0.27981013	0.9521592	0.04784082
ntp	-0.87079323	-0.3190253	0.860058	0.13994202	0.80617119	0.34585274	0.7695261	0.23047389
str	0.04729314	0.9685491	0.9403241	0.05967591	-0.06339327	-0.79207159	0.6313961	0.36860389
snt	-0.5704186	-0.7328691	0.8624745	0.13752555	0.5180198	0.78176583	0.8795023	0.12049768
Cumulative proportion of total sample variance explained	0.69210656	0.91162096			0.549	0.837		

Table 3. Comparison between pcm and ml estimations (rotated)

3. It can be concluded that a 2-factor maximum likelihood factor analysis including a vraimax rotation should be applied and the results are given in Table 4.

Variable	Estimated factor loadings		Specific variance
	F1	F2	
No. of Students per 10000	0.9746174	0.06	0.04686318
No. of Personnel per 10000	0.93480772	0.28	0.04784082
No. of Institutes per 10000	0.80617119	0.35	0.23047389
Student-Teacher Ratio	-0.06339327	-0.79	0.36860389
Senior Teachers Ratio	0.5180198	0.78	0.12049768

Table 4. Rotated factor loadings

It is clear that the first 3 variables, number of student, personnel and institutes relative to population, define factor 1, which could be called *the scale factor* measuring the scale of school running and hence the resources and opportunity available for residents in a district. Whereas, the other 2 variables load highly on factor 2, and both of them are closely related to the proficiency of education provided by teachers in a district, thus it can be named *the quality factor*.

4. Obtain factor scores for 31 observations by regression, which give measures for the scale and quality of regional higher education. Based on factor scores and the

variance contribution of each factor, composite scores are obtained as:

$$Score\_Comp = 0.69\ Score\_F1 + 0.22\ Score\_F2$$

The results are showed (ranked and sorted by composite scores) in Table 5:

Table 5. Factor Scores and ranking

District	Score_F1	Rank_F1	Score_F2	Rank_F2	Score_Comp	Rank_Comp
Beijing	3.79249953	1	1.69	2	2.989409636	1
Tianjin	1.82689095	2	-0.13	16	1.231750222	2
Shanghai	0.89997932	4	1.10	5	0.863878051	3
Shaanxi	1.27733181	3	-1.06	28	0.647128469	4
Jilin	0.56593294	5	0.18	11	0.429921161	5
Liaoning	0.41808128	7	0.31	10	0.356410411	6
Hubei	0.55261305	6	-0.13	15	0.353521933	7
Jiangsu	0.13901024	9	0.65	7	0.239172992	8
Heilongjiang	-0.12098498	14	1.17	4	0.174309324	9
Chongqing	0.39105639	8	-0.68	23	0.120715637	10
Ningxia	-0.28657148	18	0.66	6	-0.051574593	11
Fujian	-0.02513509	12	-0.23	17	-0.068515476	12
Zhejiang	-0.32289259	19	0.61	8	-0.088962517	13
Jiangxi	0.02278839	11	-0.75	25	-0.148447821	14
Shanxi	0.10331218	10	-1.28	31	-0.210495876	15
Shandong	-0.23960366	15	-0.50	22	-0.275140515	16
InnerMongolia	-0.38571061	23	-0.05	14	-0.276562108	17
Gansu	-0.41910547	24	0.00	13	-0.289200504	18
Hebei	-0.61242227	26	0.51	9	-0.309527226	19
Hainan	-0.10207799	13	-1.15	30	-0.324178513	20
Hunan	-0.3536471	21	-0.41	19	-0.333327237	21
Sichuan	-0.34120195	20	-0.68	24	-0.384789964	22
Guangdong	-0.27819941	17	-0.89	27	-0.386824793	23
Anhui	-0.25500844	16	-1.09	29	-0.416624384	24
Henan	-0.37283091	22	-0.88	26	-0.45194653	25
Guangxi	-0.59585748	25	-0.39	18	-0.497180735	26
Tibet	-1.1607406	30	1.36	3	-0.502369474	27
Qinghai	-1.61122046	31	2.73	1	-0.511024417	28
Xinjiang	-0.72090159	27	-0.41	20	-0.586914159	29
Guizhou	-0.93367261	29	0.17	12	-0.607308553	30
Yunnan	-0.85171138	28	-0.44	21	-0.68530242	31

Ⅳ. Conclusion

Via factor analysis, two factors: the scale factor and the quality factor are deprived to depict regional situation of higher education development from two perspectives.

The first factor, the scale factor, accounts for 69% of the sample variance. It measures the school-running scale and the higher education resources availability, and is the main factor on the higher education assessment. Most districts ranking top in this factor are of a excellent overall performance.

The quality factor measures the overall proficiency level of higher education. Among the cities with high composite scores, Tianjin has a relatively score on this factor, which is a strong message that future higher education development in Tianjin should shift its focus from scale to quality, by controlling the expansion in college enrollment and promoting education proficiency in existing institutes.

The composite score displays remarkable discrepancies of higher education status among districts. Beijing, Tianjin and Shanghai, which are also national and international economic centers, rank among the top in general, while the western big provinces such as Yunan, Guizhou, Xinjiang, Qinghai and Tibet are among the last. The results indicates a correlation between economic advancement and education development.

It is noteworthy that despite a low composite rank, Tibet and Qinhai are ranked among the top on quality factor which may be attributed to the large amount of financial and talent support from central government and the whole society in the course of development of western regions, while both districts has extremely low education scale due to historical, geographical and social reasons. The result suggests that the through policies and publicity, promotion of higher education coverage in those regions is vital, so as to provide more students with access to higher education, which could also promote the regional economic and cultural development.

Works Cited

Trow, M. A., & Burrage, M. (2010). *Twentieth-century higher education : elite to mass to universal*. Johns Hopkins University Press.

贾云鹏. (2008). 关于高等教育资源地域配置判断标准的几点思考. *教育科学, 24*(3), 57-61.

贾云鹏. (2009). 我国高等教育资源地域分布的现状、特点及成因. *江苏高教*(2), 48-50.

齐丹丹. (2014). 浅析我国高等教育资源地域性共享模式. *才智*(34).

王芳. *中国人口受教育程度的空间模式及其时序演变*. 华东师范大学, 2012.