Evaluation and Discussion on the Sustainable Development of Higher Education

Shichong Yang*, Yujie Dong and Tianyi Sheng

School of China Agricultural University, Beijing, China

*Corresponding author: 2018307140317@cau.edu..cn

Abstract. Though the higher education system has great significance to a nation, compelling model to evaluate the higher education system is currently lacking. Due to changes in the social environment, such as the epidemic, countries' requirements for the health and sustainability of the higher education system are gradually increasing. To assess this system, a thorough evaluation model is established by k-means cluster analysis and TOPSIS analysis. 15 indicators in 8 countries from the Global Competitiveness Report (2009-2019) are selected. In addition, we choose China to do the further analysis. Calculation results of factor analysis show that there are 5 indicators with greater influence among 15 indicators. It makes sense to increase the weights of these five indicators in China for more sustainable development.

Keywords: TOPSIS, entropy weight, higher education, sustainable system.

1. Introduction

Higher Education can benefit societies in many aspects. As education is important to the development of a nation, a right to education has been recognized by some governments and the United Nations [1-3]. The analysis has significant and far-reaching influence on the improvement and adjustment of future education in various countries. This paper establishes models which evaluate the higher education systems in different countries throng k-means cluster analysis and TOPSIS analysis, and gives suggestions for making them more sustainable based on the results.

2. Methods

2.1. Data Selection and Handling

In the analysis of the health and sustainable development of a nation's system of higher education, there are many indicators that can be used to evaluate. To build mathematic models, following indicators are collected: patent applications per million pop, tertiary education enrollment rate gross, quality of the education system, quality of math and science education, quality of management schools, Internet access in schools, availability of specialized training services, extent of staff training, capacity for innovation, quality of scientific research institutions, company spending on R&D, university-industry collaboration in R&D, government's procurement of advanced technology products, availability of scientists and engineers and intellectual property protection. When receiving the paper, we assume that the corresponding authors have granted us the copyright to use the paper in the relevant book or journal. If authors use tables or figures from other Publications, they must ask the corresponding publishers to grant them the right to publish this material in their paper.

All data is collected from The Global Competitiveness Report published annually by the World Economic Forum [4]. Data from the United States, the United Kingdom, South Korea, China, India, Brazil, Mali and Ukraine from 2009 to 2019 is used in this research. The data is chosen from different regions with different development and education environments, making it representative and comprehensive. The outbreak of a pandemic has made the data inaccurate in 2020.

In order to make statistics and calculations more convenient and efficient , the original data is processed as following:

Set parameters: rank α; total number of countries β; score θ;

Forward the ranking: θ = (β-α)/(β-1).

We make a table of the result of data processing in U.S. in Table 1.

Table 1. The Result of Data Processing in USA (parts).

Indicator	Rank	Score
Patent applications per million pop	10	0.933823529
Tertiary education enrollment rate gross	9	0.941176471
…….	……	……
Quality of scientific research institutions	5	0.970588235
Company spending on R&D	2	0.992647059
University-industry collaboration in R&D	2	0.992647059
Gov't procurement of advanced technology products	2	0.992647059
Availability of scientists and engineers	2	0.992647059
Intellectual property protection	14	0.904411765

2.2. The Establishment of Evaluate Model

2.2.1. Use entropy method to calculate weights.

As an objective weighting method, the entropy weight method can calculate the data of indicators to obtain the proportion of the given indicator in the evaluation result. Thus, the influence of different years on the evaluation of the system of higher education is obtained.

Quantify the original indicators

$$p_{ij} = \frac{x'_{ij}}{\sum_{j=1}^{n} x'_{ij}} \tag{1}$$

Calculate the information entropy of every evaluation indicator

$$e_i = -k\sum_{j=1}^{n} p_{ij}\, lnp_{ij}\ ,\ \ k = -\frac{1}{lnN} \tag{2}$$

Calculate the coefficient of variation of each indicat

$g_i = 1 - e_i$ (3)

Calculate the weight of each indicator and get the weight matrix

$$W_{ij} = \frac{g_i}{\sum_{i=1}^{m} g_i} \tag{4}$$

After the above processing, the score weight in different periods can be calculated, and the calculation formula of the comprehensive score is available according to the weight.

Taking America as an example:

S (the score of an indicator) = 0.119*Y2009 + 0.121*Y2010 + 0.134*Y2011 + 0.129*Y2012 + 0.11*Y2013 + 0.099*Y2014 + 0.117*Y2015 + 0.083*Y2016 + 0.052*Y2017 + 0.011*Y2018 + 0.022*Y2019 (5)

Using this formula, it is easy to get the comprehensive scores of the indicators of China, South Korea, India, the United Kingdom, Ukraine, the Republic of Mali and Brazil.

2.2.2. k-means Cluster Algorithm.

After getting the scores of various indicators in the time dimension, this data is put into cluster analysis to evaluate the overall situation [5].

Notations: $D = \{x_1, x_2, \cdots, x_m\}$ (input) is sample set, k is branch of clustering, N is maximum iteration and $C = \{C_1, C_2, \cdots, C_k\}$ (output) is cluster partition.

Randomly select k samples from the data set D as the initial k centroid vectors: $\{\mu_1, \mu_2, \cdots, \mu_k\}$;

For n=1, 2, ..., N

Initialize the cluster division C as $C_t = \emptyset t = 1,2,\cdots,k$;

For i=1, 2..., m, calculate the sample x_i and the distance of each centroid vector μ_j $(j = 1,2,\ldots k)$: $d_{ij} = \left\|x_i - \mu_j\right\|_2^2$, mark the smallest x_i as the category λ_i corresponding to d_{ij} and then update $C_{\lambda i} = C_{\lambda i} \cup \{x_i\}$; For j=1, 2, ..., k, recalculate the new centroid $\mu_j = \frac{1}{|C_j|}\sum_{x \in C_j} x$ for all sample points in C_j;

If all k centroid vectors have not changed, then we can go to next step;

Output cluster partition $C = \{C_1, C_2, \cdots, C_k\}$.

In the analysis, these 8 countries are divided into two categories. 6 countries have health systems and the remaining 2 countries have unhealthy systems. These countries whose system of higher education is unhealthy are have disadvantages in the education competition in the international society and their developments are full of risks.

Table 2. The Results of Comparison by Cluster Analysis(parts).

Indicator	Average Value ± Standard Deviation		F	*P*
	cluster_1(n=2)	cluster_2(n=6)		
Patent applications per million pop	0.43±0.28	0.78±0.17	4.916	0.068
Tertiary education enrollment rate gross	0.31±0.25	0.73±0.29	3.252	0.121
…….	……	……	……	……
Company spending on R&D	0.51±0.21	0.80±0.18	3.739	0.101
University-industry collaboration in R&D	0.44±0.22	0.80±0.18	5.717	0.054
Gov't procurement of advanced technology products	0.52±0.18	0.71±0.24	0.980	0.361
Availability of scientists and engineers	0.37±0.10	0.81±0.10	30.979	0.001
Intellectual property protection	0.38±0.09	0.63±0.27	1.511	0.265

Table 3. The Basic Situations of Cluster Categories.

Cluster Category	Frequency	Percentage (%)
cluster_1	2	25.00%
cluster_2	6	75.00%
sum	8	100%

2.2.3. Combine TOPSIS with entropy method.

In order to better understand the situation of higher education of each country, TPOSIS is used to analyze [6].

1) Use the entropy method to calculate the weights of different indicators

Based on the comprehensive scores of the indicators, the weights of different indicators are obtained.

2) The model of TOPSIS

a) Establish a weighted evaluation matrix

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ r_{m1} & r_{m2} & \cdots & r_{mn} \end{bmatrix} = \begin{bmatrix} W_1 x'_{11} & W_1 x'_{12} & \cdots & W_1 x'_{1n} \\ W_2 x'_{21} & W_2 x'_{22} & \cdots & W_2 x'_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ W_m x'_{m1} & W_m x'_{m2} & \cdots & W_m x'_{mn} \end{bmatrix} \tag{6}$$

b) Calculate the positive and negative ideal solution

$$V^+ = \left\{\left(\max V_{ij} \,\middle|\, j \in m\right), \mathrm{i} = 1,2,\cdots,n\right\} = \{V_1^+, V_2^+, \cdots, V_m^+\} \tag{7}$$

$$V^- = \left\{\left(\min V_{ij} \,\middle|\, j \in m\right), \mathrm{i} = 1,2,\cdots,n\right\} = \{V_1^-, V_2^-, \cdots, V_m^-\} \tag{8}$$

c) Calculate Euclidean Distance

$$D_i^+ = \sqrt{\sum_{j=1}^{m}(V_{ij} - V_j^+)} \quad (\mathrm{i} = 1,2,\cdots,m) \tag{9}$$

$$D_i^- = \sqrt{\sum_{j=1}^{m}(V_{ij} - V_j^-)} \quad (\mathrm{i} = 1,2,\cdots,m) \tag{10}$$

d) Calculate closeness

$$C_i = \frac{D_i^-}{(D_i^+ + D_i^-)} \tag{11}$$

Using the above method combined with the weight calculated by the entropy method, the comprehensive evaluation score of higher education in eight countries can be obtained. This score is reflected in the graph as C: the closeness of the relative solution. The data is closer to 1, the higher the comprehensive evaluation score of this country. In the figure, we can see that the United States has the highest score and the score indicates that its higher education system is the most complete and healthy and capable of sustainable development. However, the Republic of Mali and Brazil have the lowest scores which means that the higher education systems of the two countries are in a sub-health state and need to be improved. Combining the results with the k-means Cluster Algorithm, it can be said with certainty that they belong to the sub-health country category in the cluster analysis.

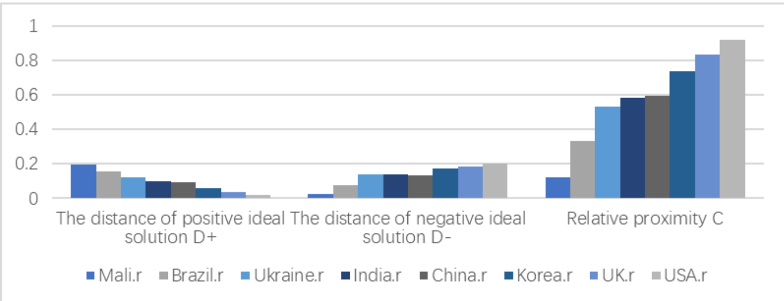

Figure 1. Entropy weight TOPSIS evaluation results.

2.3. Discussion on the Sustainable Development of China's Higher Education

Before proposing the more healthy and sustainable system of higher education, it is necessary to select several indicators with greater impact from 15 indicators first. And then we can base on these indicators to put forward more targeted measures to improve the current system of higher education. The data of these indicators is collected from the normalized data of China in the above model.

Factor analysis is one of the most important methods of multivariate statistical analysis and a statistical method that can transfer the original Variables, which affect each other, into several independent comprehensive factors. Therefore, it can compress data and also improve its quality.

2.3.1. Normalization.

We suppose that each indicator can be expressed as the following:

$$\begin{cases} x_1 = a_{11}F_1 + a_{12}F_2 + \cdots + a_{1m}F_m + \xi_1 \\ x_2 = a_{21}F_1 + a_{22}F_2 + \cdots + a_{2m}F_m + \xi_2 \\ x_3 = a_{31}F_1 + a_{32}F_2 + \cdots + a_{3m}F_m + \xi_3 \\ \quad \vdots \qquad\qquad\qquad\qquad\qquad \vdots \\ x_p = a_{p1}F_1 + a_{p2}F_2 + \cdots + a_{pm}F_m + \xi_p \end{cases} \tag{12}$$

In the formular, each variable is related to common factors so we need to determine them according to specific problems. Special factors are inconsistent with common factors.

This model needs to meet the following demands:

m<p, the number of common factors that we selected is less than the number of original variables;

var(Fi)=1, i=1,2,3,…,m;

cov (ξi, ξj)=0, var(ξi)= δi， i,j=1,2,3,…p;

cov (F, ξ)=0

The orthogonal factor model can be expressed in the form of a matrix:

$$X = AF + \xi, \tag{13}$$

$$X = (x_1, x_2, \cdots, x_p)'; \tag{14}$$

$$F = (F_1, F_2, \cdots, F_m)'; \tag{15}$$

$$\xi = (\xi_1, \xi_2, \cdots, \xi_p)'; \tag{16}$$

$$A = (a_{ij})_{p\times m}; \tag{17}$$

$$i = 1,2,\cdots, m \tag{18}$$

2.3.2. Use principal components analysis to estimate factor loading.

The normalized observation value of the original indicators in the orthogonal factor model is:

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \cdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \tag{20}$$

The eigenvalues of X are $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$, and the corresponding eigenvector are $u_1, u_2, \ldots, u_p$. Thus, the matrix of factor loading $A = \left(\sqrt{\lambda_1}u_1, \sqrt{\lambda_2}u_2, \ldots, \sqrt{\lambda_m}u_m\right)$ (m is the number of common factors that we selected). We determine the number of common factors by setting that accumulated variance contribution rate of the first m common factors ($\frac{\sum_{k=1}^{m}\lambda_k}{\sum_{i=1}^{p}\lambda_i}$) is no less than a certain threshold (e.g., 95%).

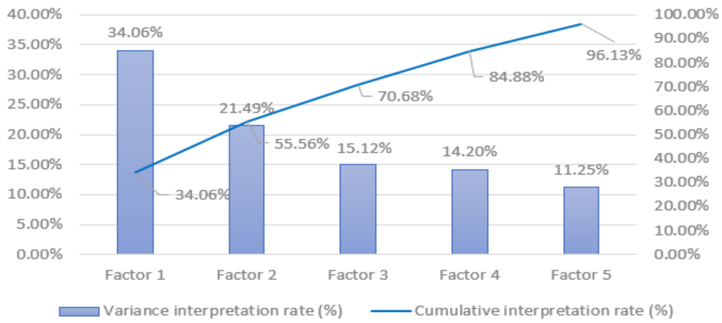

Figure 2. Factor analysis results.

2.3.3. Factor rotation.

$X = A\Gamma\Gamma' F + \xi$

After Factor analysis, we can get the weight of each indicator among the five components of factor analysis.

Among these weights, the ones that obviously have the greatest impacts on Chinese higher education are: the training of college faculty, innovation capabilities and cooperation between colleges and enterprises, the administrative efficiency of the education system and the government's procurement rate of innovative products.

We give priority to increasing the scores of these five indicators in an ideal situation, and use the model of TOPSIS to reanalyze Chinese scores. The results show that Chinese score has increased significantly under the ideal blueprint. And the scores are increase of by 20%, which verifies the correctness of our analysis.

Table 4. The Score of Original Indicators in Different Years.

Indicator	2020	2021	2022	2023	2024
Quality of the education system	0.662	0.649	0.623	0.634	0.642
Extent of staff training Extent of staff training	0.718	0.684	0.673	0.665	0.654
Capacity for innovation	0.636	0.65	0.671	0.693	0.712
Company spending on R&D	0.841	0.841	0.841	0.841	0.841
University-industry collaboration in R&D	0.771	0.795	0.799	0.785	0.776

Table 5. The Score of Ideal Indicators in Different Years.

Indicator	2020	2021	2022	2023	2024
Quality of the education system	0.87327	0.88177	0.89035	0.89901	0.90776
Extent of staff training Extent of staff training	0.87285	0.87906	0.88533	0.89163	0.89799
Capacity for innovation	0.80665	0.79438	0.78229	0.77038	0.75866
Company spending on R&D	0.8837	0.88148	0.87926	0.87705	0.87484
University-industry collaboration in R&D	0.97426	0.97664	0.97903	0.98143	0.98383

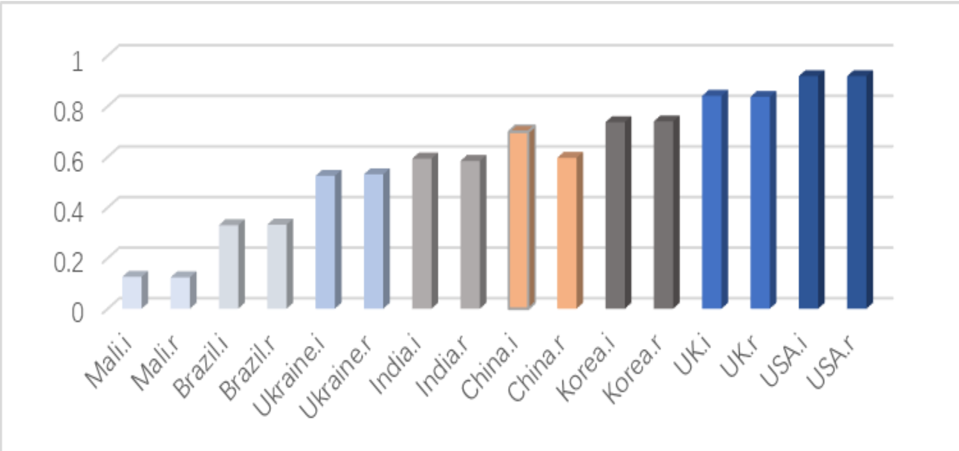

Figure 3. Comparison of relative ideal distance.

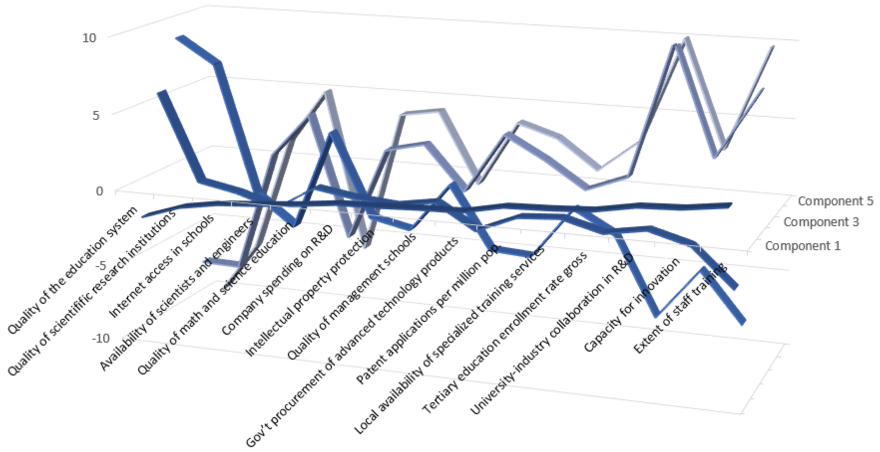

Figure 4. Score comparison of each component.

3. Result and Analysis

The results of this analysis show that those countries can be divided into two groups on the basis of the development of higher education's system. Taking some nations for examples, Mali and Brazil belong to one category, while China, the United States, India, Ukraine, the United Kingdom and

South Korea belong to another category which holds a health and sustainable system of higher education.

After classifying countries, we combine the entropy method with TOPSIS to analyze the ranking of the higher education level of different countries, based on the scoring data. In this situation, the evaluation system is established for the health level of higher education. The health of higher education in various countries can be calculated quickly and accurately through this model.

Based on the result of the score comparison, we can conclude that it is really crucial for China to cultivate college faculty, develop innovation capabilities and strengthen the cooperation between colleges and enterprises to improve the health of higher education and the ability of sustainable development. The administrative efficiency of the education system and the government's procurement rate of innovative products are secondary factors. In the further study, the five most important factors to improve the sustainable development of higher education in China are calculated. This study has reference significance for improving the quality of Chinese higher education. However, more researches are needed to determine the reform plan that best suits the national conditions.

References

[1] Article 13 of the United Nations' 1966 International Covenant on Economic, Social and Cultural Rights recognizes a universal right to education. ICESCR, Article 13.1.

[2] Francesca Pucciarelli and Andreas Kaplan. Competition and strategy in higher education: Managing complexity and uncertainty[J]. Business Horizons, 2016, 59(3) : 311-320.

[3] Wößmann, Ludger. "Why Students in Some Countries Do Better: International Evidence on the Importance of Education Policy." Education matters 1.2 (2001): 67-74.

[4] Information on https://www.weforum.org/reports

[5] Liu Qingguo et al. Multiattribute Group Decision-Making Method Using a Genetic K -Means Clustering Algorithm[J]. Mathematical Problems in Engineering, 2020, 2020.

[6] Majid Behzadian et al. A state-of the-art survey of TOPSIS applications[J]. Expert Systems With Applications, 2012, 39(17) : 13051-13069.