# **Research on Applications of Data Science in Macroeconomics**

#### Jingru Sun

College of Engineering & Applied Sciences, Stony Brook University, New York, 11794, United States of America

Keywords: Data science, Macroeconomics, Statistics

**Abstract:** Undoubtedly, it is meaningful to study China's macroeconomic system. It can not only analyze the advantages and disadvantages of the economic construction in the present region, but also provide a reference for formulating the future economic development strategy. In this paper, thirty provinces in China were selected to analyze the efficiency of Chinese macroeconomy based on BCC model of DEA. The results show that PTE is relatively high and SE is low. It is required to maintain PTE and enhance SE by the means of the development of science and technology and the expansion of investment.

#### **1. Introduction**

In recent years, China has ushered in a new wave of economic construction. All over the country, we should seize the opportunity and strive to promote economic development in the region. However, the national resources are limited, and the local resources are also limited. It involves the allocation of resources to make the limited resources play the greatest role in exchange for the greatest economic development. Especially in some areas which are supported by the state, some resources in these areas are increasing rapidly because of the inclination of the state policy. However, different regions have different investment in resources, which can bring about different economic development speed. In the region of increasing returns to scale, the output will increase with the increase of investment. In the region of diminishing returns to scale, the output of production will decrease with the increase of investment. The invalid economic system is in the region of increasing returns to scale. The reason can be understood as part of the excessive investment, or to understand that some other investment failed to keep up with the construction needs, resulting in the resources did not reach the optimal allocation. It is one of the central issues of China's sustainable development strategy to maintain the rapid, coordinated and sustainable development of the national economy is. Regional economic development plays an important role in the development of the whole national economy. Since the reform and opening up, China's economy has made rapid progress, and the people's living standards have been greatly improved. The Party Central Committee has further proposed that maintaining the sustained, healthy and rapid development of the national economy is one of the important principles for the national economic and social development in the future. Therefore, it is of great significance to deeply study the theories and methods of macroeconomic assessment. Some evaluation methods are often used in economic research. However, some methods are not reliable and objective because of the difficulty in determining the weights or the difficulty in finding explicit relationships. DEA is a non-parametric method, which has many advantages over traditional methods. From the current macroeconomic comprehensive evaluation method, DEA method is an important method to evaluate the effectiveness of economic development.

### 2. Principles and Models of DEA

#### 2.1 Principles of DEA.

Data envelopment analysis (DEA) is a new systematic analysis method developed by famous operational research scholars A. Charnes and W. Cooper on the basis of the concept of "relative

efficiency evaluation". It extends the concept of SISO engineering efficiency to the validity evaluation of decision-making units, and becomes an effective method for analyzing data characteristics. Because of the practicability of DEA method and the fact that it does not need any weight assumption, researchers in various fields soon noticed it and made it widely used in a short period of time, such as enterprise efficiency evaluation, public utilities management evaluation, provincial and municipal economic situation analysis. Especially, DEA method does not need to estimate the parameters in advance, so it has the advantages of avoiding subjective factors. simplifying algorithm and reducing errors. Data envelopment analysis is a black-box approach, which takes the weights of input-output indexes of decision-making units as optimization variables and uses data programming model to evaluate the relative effectiveness of each decision-making unit. The DEA method can effectively evaluate the relative validity of the same decision-making units under the condition. The DEA method does not need to estimate the parameters in advance, which can avoid the subjective factors given by the weight, simplify the algorithm and reduce the error. Therefore, in recent years, more and more scholars at home and abroad began to use DEA method in economic efficiency evaluation. The DEA method finds out the optimal production point in the production potential set by linear programming. The envelope surface of all optimal production points is called the effective frontier. By comparing with the effective frontier, the input redundancy or output deficiency of other production points can be calculated.

#### 2.2 CCR Model.

This paper uses DEA (Envelopment Analysis Data) to study the efficiency of Hunan logistics industry. The most basic model of DEA is the CCR model. It is assumed that the scale of remuneration is fixed to calculate the relative comprehensive technical efficiency. If there are n decision making units, m input variables for each decision unit, and p output variables.

 $X_j = (x_{1j}, x_{2j}, \dots, x_{mj})^T j = 1, 2, \dots n$  and  $Y_j = (y_{1j}, y_{2j}, \dots, y_{mj})^T j = 1, 2, \dots n$  are input variables and output variables for decision making unit DMU. CCR model is described as follows:

 $\min \theta$ 

$$s.t.\begin{cases} \sum_{j=1}^{n} \lambda_{j} \chi_{j} + s^{-} = \theta \chi_{0} \\ \sum_{j=1}^{n} \lambda_{j} \gamma_{j} - s^{+} = \gamma_{0} \\ s^{-} \ge 0, s^{+} \ge 0, \lambda_{j} \ge 0, \\ \theta \end{cases}$$

#### 2.3 BCC Model.

In terms of evaluation methods, the existing literature evaluation methods are all based on the CCR model of DEA. CCR model can only distinguish the economic efficiency units from the ineffective ones and cannot support the grouping and full sorting of economic efficiency units, and cannot examine the impact of random errors on economic efficiency. Therefore, the evaluation results obtained by this method are general and complex systems. The study of evaluation methods is of great significance to economic and social development, but it is also a very arduous work. In the application of DEA method to evaluate complex systems, this method not only has unique advantages, but also has shortcomings. It can not only supplement and perfect the existing evaluation methods of complex systems, but also have positive significance to the evaluation of regional economic systems. Depending on whether scale reward is changeable or not, DEA model can be divided into CCR model with constant scale reward and BCC model with changeable scale reward. According to the hypothesis of the two models, if all the decision-making units operate at the optimal scale, the CR model with constant return on scale can be selected. But in fact, China's financial development has

been constrained by the economy and finance, so there is incomplete competition. It is impossible for all listed forestry companies to operate at the optimal scale. BCC model considers both technical efficiency and scale efficiency.

## 3. Empirical Research of Macroeconomy Efficiency Based on DEA

Macro economy is a complex system with multiple inputs and multiple outputs. The evaluation index should be considered to achieve the purpose of evaluation. That is to say, the selection of input and output indicators should serve and obey the purpose of systematic evaluation: the index system should fully reflect the purpose of evaluation. Indicators that have a greater impact on the purpose of the evaluation should be included; the links between input and output indicators should be considered. When there is a strong correlation between a certain index and other indicators, it can be considered that the information of the index has been included to a large extent by other indicators, so it is not necessary to regard it as an input-live-output index. When the indicators number is too small, it is not conducive to find out the causes of poor economic operation. If we want to refine the index system, it may be affected by the ambiguity of the system boundary. Using DEA method to study economic efficiency, firstly, input and output indicators should be selected reasonably and scientifically based on the purpose of system evaluation, and the principle of DEA method should be satisfied: the number of indicators is at most half of the number of samples or the product of the number of input and output indicators is not greater than the number of samples. Based on the neo-classical growth theory that output growth is mainly driven by labor factors of capital and technological progress in economics, this paper chooses input and output indicators according to the characteristics of China's economic development and the availability of data. Among them, the input index reflects the proportion of fixed assets investment in GDP, the number of employment personnel, the investment in education and the environmental protection expenditure of green economy and ecological environment. The corresponding output indicators mainly reflect the comprehensive strength of the provinces and cities of GDP, reflect the average per capita GDP and reflect the level of the ecological environment of the three wastes comprehensive utilization of product output value in GDP. The selection of the above-mentioned indicators mainly highlights the level of economic efficiency of the provinces, and comparative analysis of the economic efficiency of the provinces has certain practical significance. We use the BCC model to get the result of operation.

## 4. Results Analysis and Promotion Paths of Macroeconomy Efficiency

## 4.1 Technology Efficiency.

The comprehensive efficiency (also called technology efficiency) of Beijing, Shanghai and Zhejiang has reached a peak of 1. The minimum value is only 0.520, which belongs to Jiangxi province. The low technical efficiency of macroeconomy is mainly caused by two aspects: pure technical efficiency and scale efficiency.

## 4.2 Pure Technology Efficiency.

Eight provinces' pure technology efficiencies are below 0.8, which are Xinjiang, Guangxi, Hainan, Ningxia, Liaoning, Jilin, Gansu, Guangdong and Jiangxi. We should develop advanced science and technology and adhere to independent innovation in science and technology and institutional innovation in development. The above analysis shows that China's overall economic efficiency is good, but it has not reached the optimal state of the whole, and its development space is vast. Therefore, we can improve our economic efficiency, deepen the openness of regional economy, and promote cooperation and exchanges among regions by deepening self-innovation in science and technology and combining institutional and management innovation. From the above analysis, we can see that the supply and demand of economic resources are unbalanced in different regions, some provinces have redundant input of resources and waste of resources, while the resources of the

original provinces are insufficient. Therefore, the inter-regional flow of input factors can be promoted through the opening and cooperation between provinces and cities to improve the utilization of resources

Province	TE	PTE	SE	SR
Anhui	0.733	0.888	0.825	irs
Beijing	1.000	1.000	1.000	-
Fujian	0.778	0.927	0.839	irs
Gansu	0.645	0.725	0.889	irs
Guangdong	0.595	0.735	0.809	irs
Guangxi	0.682	0.788	0.866	irs
Guizhou	0.647	0.893	0.725	irs
Hainan	0.768	0.771	0.996	irs
Hebei	0.734	0.969	0.757	irs
Heilongjiang	0.688	0.912	0.754	irs
Henan	0.835	0.969	0.862	irs
Hubei	0.742	0.855	0.868	irs
Hunan	0.733	0.812	0.903	irs
Inner Mongolia	0.728	0.886	0.822	drs
Jiangsu	0.855	0.988	0.865	irs
Jiangxi	0.520	0.605	0.860	irs
Jilin	0.636	0.732	0.869	irs
Liaoning	0.632	0.742	0.852	irs
Ningxia	0.639	0.753	0.848	irs
Qinghai	0.720	0.972	0.741	irs
Shaanxi	0.772	0.819	0.943	drs
Shandong	0.744	0.845	0.881	irs
Shanghai	1.000	1.000	1.000	-
Shanxi	0.762	1.000	0.762	irs
Sichuan	0.796	0.889	0.895	irs
Tianjing	0.735	0.821	0.895	irs
Tibet	0.799	0.899	0.889	irs
Xinjiang	0.787	0.796	0.989	irs
Yunnan	0.982	1.000	0.982	irs
Zhejiang	1.000	1.000	1.000	-

Table 1. Macroeconomy efficiency of all provinces in China

### 4.3 Scale Efficiency and Scale Return.

From the perspective of scale efficiency, except Beijing, Shanghai and Zhejiang, the other provinces are below 1. Especially, the scale efficiencies of Hubei, Heilongjiang, Inner Mongolia, Qinghai and Guizhou are below 0.8. Except Shaanxi and Inner Mongolia, the scale income of other provinces and cities is less than 1, in the increasing range of scale income, and the scale income value is relatively small. According to the scale income theory, if we increase capital investment, we will be able to exchange for greater output. This shows a problem: all provinces and cities in China are basically in a state of shortage of funds. The problem of capital is the bottleneck of the development

of our province's macro-economy. China's provinces and cities are in a state of capital shortage, looking for funds is one of the urgent problems to be solved in the development of the secondary industry economy. When solving capital problems, we should focus on promoting private capital and increasing the proportion of self-financing. Gradually reduce the state budgetary capital investment and save national resources. When considering the capital investment, we should not blindly invest in infrastructure, nor blindly increase the investment in renovation and transformation. We should carefully find a balance between the two investments. According to the value of scale benefit, each region can formulate the investment plan for the next year. If the scale benefit is less than 1, the output can be increased by expanding input. If the return on scale is 1, the investment should be fully absorbed and digested from the perspective of improving productivity.

### References

[1] Zhang Tianhua, Chen Li, Dong Zhiqiang. Highway Construction, Firm Dynamics and Regional Economic Efficiency [J]. China Industrial Economics, 2018(1): 79-99.

[2] Wang Yizhong, Frank M. Song. Macroeconomic Uncertainty Demand for Financing and Corporate Investment [J]. Economic Research Journal, 2014(2): 4-17.

[3] Fieldhouse A J, Mertens K, Ravn M O. The macroeconomic effects of government asset purchases: evidence from postwar US housing credit policy[J]. The Quarterly Journal of Economics, 2018, 133(3): 1503-1560.

[4] Ben Zeev N, Pappa E. Chronicle of a war foretold: The macroeconomic effects of anticipated defence spending shocks[J]. The Economic Journal, 2017, 127(603): 1568-1597.