Dynamically Evaluating Innovation Efficiency and Its Determinants of Public Research and Development (R&D) Systems

Qianqian Yuan(✉)

School of Economics and Management, Shanghai Institute of Technology, Shanghai, China
1565492092@qq.com

Abstract. Public Research and Development (R&D) assumes an important function of basic knowledge and common technology supply in the R&D system and has obvious public product characteristics, which is beneficial to the knowledge stock of the whole society. In this paper, we evaluate and study the public R&D output efficiency, its changing trend and influencing factors of 30 provinces in China from 2010 to 2020 by using DEA window model and Tobit model, which is beneficial to policy makers to optimize resource allocation and promote high-quality regional economic development.

Keywords: Public R&D · DEA · Scientific and technological innovation · Efficiency evaluation

1 Introduction

Public R&D, as the foundation of science and technology innovation, is a science and technology R&D activity that is financially supported and led by the government, with public R&D institutions and higher education institutions as the main R&D subjects, mainly for basic and applied research. For a long time, governments and scholars have regarded public R&D investment as an important means to enhance national innovation capacity and national competitiveness and to promote sustainable and stable economic and social development [1].

According to the report of China Science and Technology Statistical Yearbook, it is known that China invested a total of 2,214.36 billion yuan in research and development (R&D) in 2020, an increase of 246.57 billion yuan over the previous year, with an annual growth rate of 12.5%. Since 2006, China has gradually developed into the second largest R&D investment country after the United States, accounting for nearly 30% of the world's R&D investment. Both in the scale of R&D investment and R&D intensity have been ranked among the world's leaders. However, public R&D investment accounts for only 17.3%, with an annual growth rate of about 0.7%. In the same period, the United States and Japan public research funding accounted for more than 30%, the United Kingdom and France is as high as 60%, which is much lower than the proportion of public R&D funding in developed countries.

V. Balakrishnan et al. (Eds.): ICMEIM 2022, AHSSEH 6, pp. 641–649, 2022.
https://doi.org/10.2991/978-94-6463-044-2_82

Most of the previous studies are static efficiency evaluations based on traditional DEA models, and relatively few studies on dynamic efficiency evaluation of public R&D, which has limitations in explaining efficiency trends and causes of changes. In this paper, we use inter-provincial panel data from 2010 to 2020 to measure public R&D output efficiency, present the dynamic evolution process of output efficiency by applying DEA window analysis, and analyze its influencing factors.

2 Methodology

2.1 DEA Window Analysis

Data envelopment analysis (DEA) has the advantages of not assuming the functional form of the input-output relationship in advance, avoiding the interference of subjective factors and being able to measure the efficiency of multiple inputs and outputs simultaneously, but the traditional DEA can only analyze the data statically, and the result cannot be compared across periods to observe their dynamic trends. To address this shortcoming, Charnes et al. first proposed the DEA window analysis method in 1985 [2], which not only reuses the number of decision units and increases the sample size, but also treats different periods of the same decision unit as different decision units, and thus can obtain a more realistic efficiency evaluation.

Suppose that the efficiency of N DMU decision units in time period T needs to be evaluated, and each DMU has r inputs and s outputs, then N $\times$ T sample observations need to be observed. For a particular decision unit $\mathrm{DMU_n}$ in period t, the r-dimensional input vector is $X_t^n = \left(x_{1t}^n, x_{2t}^n, \ldots, x_{rt}^n\right)$ and the s-dimensional output vector is $Y_t^n = \left(y_{1t}^n, y_{2t}^n, \ldots, y_{rt}^n\right)$, , where n = 1, 2, … , N. Assuming that the window time starts from k ($1 \leq k \leq T$) and the window width is w ($1 \leq w \leq T - k$), there are n $\times$ w DMUs. The input-output matrix for the kth window is shown below.

$$X_{kw} = \left(x_1^k, x_2^k, \ldots, x_N^k, x_1^{k+1}, x_2^{k+1}, \ldots, x_N^{k+1}, \ldots, x_1^{k+w}, x_2^{k+w}, \ldots, x_N^{k+w}\right) \quad (1)$$

$$Y_{kw} = \left(y_1^k, y_2^k, \ldots, y_N^k, y_1^{k+1}, y_2^{k+1}, \ldots, y_N^{k+1}, \ldots, y_1^{k+w}, y_2^{k+w}, \ldots, y_N^{k+w}\right) \quad (2)$$

Applying the output-oriented BCC model within the kth window gives the analytical results of the DEA window analysis model, and the output efficiency value for the nth DMU in period t for the kth window can be obtained by solving the following linear programming [3].

$$\begin{gathered}\mathrm{Min}\theta_0 = \theta_n^k(X_{n0}^k, Y_{n0}^k) \\ \text{s.t.}\begin{cases}\sum_{n=1}^{N} \lambda_n^k X_{kw} \leq \theta_0 X_{n0} \\ \sum_{n=1}^{N} \lambda_n^k Y_{kw} \geq Y_{n0} \\ \sum_{n=1}^{N} \lambda_n^k = 1, \lambda_n^k \geq 0 \\ 1 \leq t \leq T, 1 \leq n \leq N\end{cases}\end{gathered} \quad (3)$$

Since this paper examines the research interval of 20010–2020, considering the lag of science and technology output results, the DEA window analysis model with a lag period of one year is used, and according to the relevant research of previous scholars, using a relatively small window width can make the obtained efficiency values more accurate, so the window width w = 5 is chosen here.

2.2 Tobit Model

Since the public R&D output efficiency values calculated by DEA method are all between 0 and 1, if the traditional OLS method is used to estimate the degree of influence of each environmental factor on output efficiency, bias and inconsistency may occur. Therefore, in this paper, a Tobit model dealing with limiting variables is used to analyze the influencing factors of public R&D output efficiency, and a panel data model is constructed in the following specific form.

$$\mathrm{Y}_{\mathrm{it}}^{*} = \beta_0 + \beta_{\mathrm{i}}\mathrm{X}_{\mathrm{it}} + \varepsilon_{\mathrm{it}}$$

$$\begin{cases} Y_{it} = \mathrm{Y}_{\mathrm{it}}^{*}\mathrm{Y}_{\mathrm{it}}^{*} > 0 \\ Y_{it} = 0\mathrm{Y}_{\mathrm{it}}^{*} \leq 0 \end{cases} \tag{4}$$

where $\mathrm{Y}_{\mathrm{it}}^{*}$ is the efficiency of public R&D output in each province, X_{it} is the independent variable, β_0 is the intercept term, β_{it} is the regression coefficient of the independent variable, and $\varepsilon_{\mathrm{it}}$ is the random error term.

3 Variables and Data

3.1 Input and Output Variables

Considering that public R&D activities are a complex process of multidimensional inputs and multidimensional outputs, this paper selects the internal expenditure of R&D funds of universities and public research institutions as the index of research capital input, the full-time equivalent of R&D personnel as the index of research manpower input; and the number of published scientific and technical papers, the number of patent applications, the number of published scientific and technical works and the transfer or licensing income of patent ownership of universities and public research institutions in each province as the index of scientific and technical output, these six indexes constitute the evaluation system of public R&D input and output indexes (Table 1).

Table 1. Public R&D innovation efficiency evaluation index system [Owner-draw]

Indicator Type	Indicator Name	unit
Input indexes	Internal expenditure of R&D funds	10,000 Yuan
	Full time equivalent of R&D personnel	man-year
Output indexes	Number of scientific and technical papers	piece
	Number of patent applications	piece
	Number of scientific and technical publications	piece
	Patent ownership transfer or licensing income	10,000 Yuan

Table 2. Factors influencing the efficiency of public R&D output [Owner-draw]

Measurements	Symbols	Indicator Description
Informatization level	infor	Regional Internet access as a proportion of total regional population (%)
R&D Staff Quality	edu	The ratio of master's and doctoral researchers to total researchers (%)
R&D Network Size	size	Number of regional universities and public research institutions (pcs)
Intellectual property Protection efforts	ipr	Regional technology market transactions as a proportion of regional GDP (%)
Government R&D Support	gov	Natural logarithm of government subsidies for public R&D funding

3.2 Selection of Influencing Factors

Through a review of previous literature, this paper considers the influencing factors affecting the efficiency of regional public R&D output in five aspects: (1) Information level (infor): since the official comprehensive index has not been disclosed, and the measurement faces certain difficulties and challenges, the Internet penetration rate is adopted here as a measure of a region's informatization level [3]. (2) The quality of R&D personnel (edu) is selected as a measure of the ratio of master and doctoral researchers to total researchers, and R&D personnel quality reflects the ability to accept and master new knowledge and technology [4]. (3) The size of R&D network (size) facilitates the alignment behavior of R&D individuals and knowledge flow, thus positively influencing the R&D output performance of a region [3]. Therefore, the number of R&D organizations is chosen to measure the size of regional R&D networks [5]. (4) The strength of intellectual property protection (ipr) can stimulate R&D organizations to create new knowledge and technology and provide a good science and technology base. In this paper, the ratio of total technology market turnover to GDP is used as a measure of the strength of IPR protection in a region. (5) Government R&D support (gov): In order to examine the influence of government financial support on the efficiency of R&D output at this stage, this paper selects the natural logarithm of government R&D subsidies to public R&D organizations as a measure (Table 2).

The data of the above variables are mainly obtained from the China Statistical Yearbook, China Science and Technology Statistical Yearbook and China Internet Development Statistical Report released by the National Bureau of Statistics from 2010 to 2021, and some missing data are interpolated to make up for the missing data.

4 Empirical Analysis

4.1 Results of Public R&D Innovation Efficiency

Based on the window analysis model, this paper uses Max-DEA software, mainly analyzes the DEA-BCC model with a window width of 5 to dynamically evaluate the output

Table 3. Innovation efficiency of public R&D systems [Owner-draw]

Province	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	Average
Beijing	1.000	1.000	1.000	1.000	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Tianjin	0.619	0.558	0.508	0.554	0.546	0.577	0.523	0.540	0.587	0.662	0.534	0.564
Hebei	0.786	0.764	0.791	0.684	0.664	0.659	0.694	0.702	0.596	0.626	0.592	0.687
Shandong	1.000	0.903	0.839	0.816	0.886	0.910	0.832	1.000	0.969	0.989	0.996	0.922
Jiangsu	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.997	1.000	1.000	1.000	1.000
Shanghai	1.000	0.856	0.835	0.737	0.744	0.802	0.784	0.782	0.871	0.865	0.783	0.824
Zhejiang	1.000	0.993	0.989	0.950	0.919	1.000	1.000	1.000	1.000	1.000	0.860	0.974
Fujian	1.000	0.966	0.992	0.919	0.804	0.914	0.919	0.821	0.733	0.676	0.554	0.845
Guangdong	1.000	0.994	0.957	0.876	0.843	0.942	0.989	0.910	0.990	1.000	0.957	0.951
Hainan	1.000	0.970	0.911	1.000	0.996	1.000	1.000	0.819	0.701	0.627	0.926	0.905
Liaoning	1.000	1.000	0.883	0.796	0.766	0.863	0.842	0.801	0.732	0.700	0.639	0.820
East	0.946	0.909	0.882	0.848	0.833	0.879	0.871	0.852	0.834	0.831	0.804	0.863
Shanxi	0.688	0.677	0.592	0.554	0.484	0.600	0.652	0.723	0.825	0.712	0.724	0.657
Anhui	0.842	0.750	0.778	0.673	0.656	0.714	0.716	0.795	0.742	0.751	0.698	0.738
Jiangxi	0.916	0.893	0.960	0.899	0.874	0.880	0.935	1.000	0.930	0.910	0.828	0.911
Henan	1.000	1.000	1.000	0.942	0.949	1.000	0.989	1.000	1.000	0.964	0.883	0.975
Hubei	1.000	0.972	0.987	0.936	0.978	0.933	0.964	0.999	1.000	0.981	0.857	0.964
Hunan	0.973	0.948	0.859	0.861	0.889	0.969	0.952	1.000	1.000	0.986	0.626	0.933
Neimenggu	0.772	0.854	0.781	0.810	0.668	1.000	1.000	0.912	0.836	0.756	0.710	0.827
Guangxi	1.000	0.942	0.974	0.814	0.953	0.990	0.929	0.795	0.649	0.679	0.765	0.863
Jilin	0.672	0.631	0.578	0.565	0.563	0.655	0.709	0.696	0.682	0.707	0.618	0.643
Heilongjiang	0.665	0.786	0.835	0.702	0.642	0.717	0.644	0.708	0.753	0.779	0.674	0.719
Middle	0.853	0.845	0.834	0.776	0.766	0.846	0.849	0.863	0.842	0.823	0.758	0.823
Chongqing	0.968	0.913	0.976	0.905	0.875	1.000	0.945	0.921	0.876	0.747	0.680	0.892
Sichuan	0.602	0.606	0.580	0.552	0.560	0.589	0.576	0.591	0.608	0.582	0.574	0.584
Guizhou	0.979	1.000	0.989	0.966	0.892	0.912	0.929	1.000	0.824	0.880	0.857	0.929
Yunnan	0.694	0.756	0.793	0.678	0.747	0.825	0.712	0.602	0.647	0.641	0.570	0.697
Shaanxi	0.519	0.557	0.532	0.494	0.480	0.543	0.586	0.630	0.671	0.685	0.629	0.575
Gansu	0.864	0.672	0.771	0.685	0.711	0.753	0.762	0.704	0.672	0.660	0.534	0.703
Qinghai	1.000	1.000	0.887	0.927	0.938	1.000	1.000	1.000	0.843	1.000	1.000	0.963
Ningxia	1.000	1.000	0.939	0.888	0.974	0.735	0.920	0.899	0.675	0.757	0.885	0.879
Xinjiang	0.967	0.978	0.896	0.827	0.795	0.931	0.924	0.785	0.804	0.684	0.771	0.851
West	0.844	0.831	0.818	0.769	0.775	0.810	0.817	0.792	0.736	0.737	0.722	0.786
National	0.884	0.865	0.847	0.800	0.793	0.847	0.848	0.849	0.807	0.800	0.764	0.827

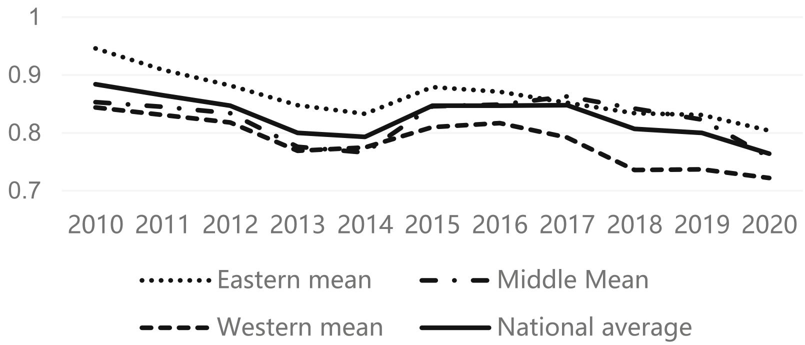

Fig. 1. Average output efficiency values for each region [Owner-draw]

efficiency of public R&D in each province in China from 2010 to 2020, and the results are shown in Table 3.

Table 3 gives the output efficiency values of public R&D by region for 2010–2020. Overall, the national average value of public R&D output efficiency for 2010–2020 is 0.827, with the maximum value of efficiency as high as 1.000 and the minimum value as low as 0.480, and there is still much room for improving the output efficiency of each province. Comparing the efficiency values of the three regions, it can be seen that the average output efficiency of public R&D in the eastern region is at the highest level in the country during the sample period. In specific analysis, the average output efficiency of the eastern, central, and western regions are 0.863, 0.823, and 0.786, respectively, while the central and western regions do not reach the average efficiency level of the country.

In terms of dynamic trends, the public R&D output efficiency of each province does not show a significant growth trend over time, but rather shows signs of regression. Figure 1 portrays the differences between the average output efficiencies obtained by each region based on the model. By and large, the eastern region has higher average efficiency values than the leading region, but during 2017–2018, the central region caught up with the eastern region to become the most efficient region in terms of output efficiency, and the western region has been lagging behind more. The trend line of the national average R&D efficiency value shows that the overall output efficiency shows a wave trend of falling, then rising, then falling again (Fig. 1).

4.2 Analysis Determinants of Public R&D Innovation Efficiency

The descriptive statistics of the influencing factors are shown in Table 4. As can be seen from the table, the east regional mean value is higher than others, indicating that the east has an advantage in all the influencing factors of public R&D output results, followed by the central and western regions in that order. In terms of intellectual property protection, the eastern region is ahead of the western region and the central region is the lowest.

To avoid the existence of multicollinearity among variables, the correlation among the main variables was analyzed in this paper, and the variance inflation factor (VIF)

Table 4. Descriptive statistics for each variable by region in China [Owner-draw]

Variable Name	National		East		Middle		West	
	mean	sd	mean	sd	mean	sd	mean	sd
infor	46.89	13.94	57.60	12.33	40.72	10.82	40.67	10.63
edu	0.57	0.09	0.61	0.07	0.57	0.09	0.54	0.09
size	205.32	94.21	244.07	115.34	220.04	42.00	141.61	73.10
ipr	1.28	2.48	2.12	3.75	0.61	0.74	0.99	1.16
gov	12.70	1.29	13.36	1.24	12.54	0.69	12.09	1.50

Table 5. Empirical results of factors influencing the innovation efficiency of public R&D [Owner-draw]

Variable/Region	National	East	Middle	West
infor	0.003**	-0.004	0.007***	0.004**
	(2.49)	(-1.62)	(3.64)	(1.89)
edu	0.264*	0.872***	0.272	-0.081
	(1.78)	(2.79)	(1.12)	(-0.36)
size	0.001***	0.001***	0.005*	0.001**
	(4.89)	(4.02)	(1.38)	(2.29)
ipr	0.024***	0.021**	0.009	0.031***
	(3.45)	(2.04)	(0.67)	(2.96)
gov	-0.191***	-0.148***	-0.230***	-0.171***
	(-10.23)	(-4.31)	(-7.90)	(1.50)
Cons	2.733***	2.191***	4.04***	2.55***
	(13.68)	(6.35)	(9.58)	(10.99)
Observations	330	121	110	99
Log likelihood	242.212	54.230	96.672	113.255
Prob > chi2	0.000	0.000	0.000	0.000

Note: t-statistics in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

method was used for multicollinearity diagnosis. The results show that there is no serious multicollinearity in the model and the variables are appropriately selected.

According to the constructed Tobit model of public R&D innovation efficiency, regressions were conducted using Stata15.1, and the regression results were estimated for the eastern, central, western, and northeastern regions of China respectively, considering the differences in the factors influencing public R&D output efficiency in each region of China, and the regression results were obtained as shown in Table 5.

From the regression results of the Tobit model in Table 5, the level of regional informatization (infor) has a significant positive effect on the innovation efficiency of public R&D (except for the eastern region), which indicates that the higher the level of regional informatization, the higher the ability of information transfer and acquisition, and the more conducive to improving regional research efficiency. The quality of R&D personnel (edu) has a positive correlation with innovation efficiency in the eastern region, with a t-test value significant at 1%, which is consistent with the regression results of the total national sample, indicating that public R&D has higher requirements for workforce quality, and the higher the quality of regional R&D personnel, the better it can make full use of available resources for R&D innovation, which in turn improves the efficiency of R&D output. The size of public R&D network (size) has a significant positive effect on the efficiency of R&D output, and the variability between regions is small. The strength of intellectual property protection (ipr) shows a positive effect on innovation efficiency in all regions of the country, but not significantly in the central region, indicating that to a certain extent, intellectual property protection can stimulate R&D organizations to research and develop new knowledge and technology and enhance regional innovation efficiency. Government R&D funding (gov) has a significant negative effect on the efficiency of public R&D output in each region. Government funding support plays an important role in improving the public R&D environment, but the positive effect is not apparent at this stage, which may indicate that the current governmental public R&D investment in China is inefficient and needs to be targeted to support R&D projects and adjust the investment structure.

5 Conclusion

This paper uses DEA window model and panel data model to analyze the current situation of innovation efficiency of public R&D system and regional differences of its influencing factors using data related to public R&D inputs and outputs from 2010–2020, and finds that (1) in terms of public R&D innovation efficiency, the eastern region is significantly higher than the central and western regions, and the gap of innovation efficiency values in the western region is widening, and the polarization phenomenon serious. (2) In terms of factors influencing public R&D innovation efficiency, informationization level, R&D personnel quality, R&D network scale and intellectual property protection have positive promotion effect on innovation efficiency, while government R&D support has negative effect on output efficiency instead. Therefore, while creating a better R&D environment for public R&D system, we should also focus on the reasonable use of government funds.

References

1. Sujin Min, Juseong Kim, Yeong-Wha Sawng. The effect of innovation network size and public R&D investment on regional innovation efficiency[J]. Technological Forecasting & Social Change, 2020, 155: 1-13.
2. Charnes A, Clark CT, Cooper WW, et al. A developmental study of data envelopment analysis in measuring the efficiency of maintenance units in the U.S. air forces[J]. Annals of Operations Research, 1985, 2(1):95-112.

3. Shoufu Lin, Ruoyun Lin, Ji Sun, Fei Wang and Weixiang Wu. Dynamically evaluating technological innovation efficiency of high-tech industry in China: Provincial, regional and industrial perspective[J]. Socio-Economic Planning Sciences 2020: 1-20.
4. Kaihua Chen and Jiancheng Guan. Measuring the Efficiency of China's Regional Innovation Systems: Application of Network Data Envelopment Analysis (DEA)[J]. Regional Studies, 2014(46): 355-377.
5. Ronald, S., and Burt. The network structure of social capital[J]. Research in Organization Behavior, 2000(22):345-423.