International Journal of Business Marketing and Management (IJBMM)
Volume 8 Issue 4 July-Aug 2023, P.P. 24-33
ISSN: 2456-4559
www.ijbmm.com

Evaluation of Innovation Efficiency and Its Influencing Factors: An Empirical Analysis of Chinese Big Data Listed Companies

Zhenzhu Wang, Qing He
School of Economics and Management, Harbin Institute of Technology (Weihai), China

Abstract : *Big data has become the core element of ubiquitous empowerment and integrated innovation of enterprises. The innovation efficiency of big data companies is the key to promote the innovation and development of digital economy. Based on the data of 55 big data listed companies in China from 2016 to 2020, this paper used data envelopment analysis and Malmquist index analysis to measure the innovation efficiency of different big data companies, and evaluated the technical innovation efficiency level and relative development trend of China's big data industry from both static and dynamic aspects. The results show that the technological innovation efficiency of big data companies in China is very important to improve their comprehensive innovation efficiency. Through Tobit regression, it is further analyzed that the social capital of top management team is an important factor to improve the innovation efficiency of companies. To this end, the government should further provide positive policy conditions for the innovation and development of China's big data industry. Managers should further optimize the allocation of resources, accelerate the introduction of innovative talents, improve and maintain the social capital of senior management teams, and inject strong impetus into the innovation of China's big data industry.*

Keywords - *Big Data Listed Companies, Data Envelopment Analysis, Innovation Efficiency, Malmquist Index Analysis, Tobit Regression*

I. INTRODUCTION

With the rapid development of cloud computing, Internet of Things and computer information technology industries, big data has become an important strategic resource of the country. The digital economy has become a new kinetic energy for economic growth in the world today. Big data is an important pillar for the transformation of old and new economic kinetic energy. During the "Thirteenth Five-Year Plan" period, China's big data industry also started rapidly and achieved rapid results, growing into an advantageous industry that drives China's economic development and social change. During the "Fourteenth Five-Year Plan" period, the rapid development of big data industry started the transformation of digital economy and society. The report of the 19th National Congress of the Communist Party of China clearly stated that "innovation is the first driving force for development and the strategic support for building a modern economic system". Enterprise technological innovation is the key factor to promote the sustainable development of enterprises. The remarkable efficiency of technological innovation not only improves the business performance of enterprises, but also effectively promotes the allocation of resources. The technological innovation of big data enterprises plays a vital role in the transformation from real economy to digital economy. The efficiency of technological innovation is the core index to measure the performance level of technological innovation of enterprises.

With the "blowout" growth of data, a large number of application-oriented big data enterprises such as data processing, data classification and data application have emerged. Big data has become the most valuable digital wealth of an enterprise and also plays a vital role in the external business circulation of enterprises.

Most of scholars' research on big data enterprises focuses on the application of big data. There are applications of big data in the spatio-temporal analysis of Seoul subway traffic [1] and medical diagnosis [2]. Alice Huang(2019) applied big data to the publishing industry to better provide knowledge services [3]. There are also studies that provide a data collection-development matrix for the government with weak big data application ability, which is used to improve the ability of data collection and data development and help the government solve social problems [4]. John and Eduardo(2016)[5] believe that analysis and big data will consume resources. If the big data application process is ineffective, it will cause significant strategic risks. The technological changes brought by big data analysis are changing the way we collect and view data. To sum up, the literature on big data focuses on the development and prediction of big data industry and the practical application of big data, and big data has become the mainstream of social development.

The research on innovation efficiency mainly focuses on enterprises. Ferrier(2017)[6] takes 1074 American surgical hospitals from 2011 to 2016 as the research object, and uses the two-stage DEA model to measure and evaluate the innovation efficiency; Justin (2019) [7] and others explored the role of internal R&D and external knowledge in the innovation of small and medium-sized enterprises in Ireland from the perspective of measuring innovation efficiency through certain research data. Claudio et al. (2013)[8] put forward an innovative management mode based on input-output mode, and expounded the importance of technological innovation efficiency of enterprises from the perspective of company managers; Kang et al. (2012) [9] combined the concept of national innovation system with innovation value chain, and found that both government subsidies and enterprise R&D cooperation have a positive impact on the innovation efficiency of Korean biotechnology enterprises; Taghizadeh et al. (2017) [10] divide the innovation activities of enterprises into three stages, and on this basis, study the influence of innovation strategies on innovation efficiency.

By combing the relevant literature at home and abroad, it is found that scholars from all over the world mainly focus on the definition of big data, the analysis and application of big data in different industries, the technology and application of big data mining, etc., and there is little research on big data enterprises themselves, and even less research on the technological innovation performance of big data enterprises.

This paper takes the listed big data enterprises in China as the research object, and establishes a reliable evaluation system of technological innovation efficiency by selecting appropriate input variables and output variables. The BBC model in Data Envelopment Analysis (DEA) is used to measure and evaluate the technological innovation efficiency in each stage. In addition, Tobit regression model is used to analyze the influencing factors of technological innovation in big data enterprises, so as to explore the "black box" in the innovation process of big data enterprises, and put forward relevant policies and suggestions for the healthy development and effective improvement of technological innovation efficiency of big data enterprises in China. It is also hoped that this empirical study can provide some reference value for technological innovation of enterprises in big data industry.

II. RESEARCH DESIGN

2.1 DATA SOURCES

By 2022, there were 266 A-share and B-share big data concepts, including 79 listed companies with big data concepts on the main board. On this basis, listed companies with incomplete data between *ST, ST and 2016-2020 were excluded. This paper finally collected the data of 55 big data listed companies from 2016 to 2020. The data mainly came from the company's annual report and CSMAR database.

2.2 METHODS

Data Envelopment Analysis (DEA). Select appropriate innovation efficiency measurement variables, with R&D personnel, R&D funds and fixed assets as input variables, patents, total profits, intangible assets and doctoral dissertations as output variables, and years of establishment, enterprise scale, economic development level and equity nature as environmental variables. The data envelopment model (DEA-BCC) with changing returns to scale is used to measure the technological innovation efficiency of big data enterprises. On the basis of static analysis, the dynamic change of technological innovation efficiency is measured by Malmquist index model.

Tobit regression model. Because the efficiency value estimated by DEA-BCC is in the range of 0 to 1, and it has truncation characteristics. In this paper, Tobit model is used for regression analysis to explore the important influencing factors. Using Tobit model to modify the estimation results of DEA model can prevent significant deviation in estimation and improve the accuracy of model estimation.

2.3 MEASUREMENT OF ENTERPRISE INNOVATION EFFICIENCY

Nasierowsli and Arcelus(2003)[11] made an early evaluation of the technological efficiency of enterprise innovation. In terms of innovation input indicators, they think that it is more appropriate to choose the amount of imported goods, R&D expenditure and education expenditure as input indicators, and the number of patent applications, productivity and per capita GDP as output indicators. Elias G.Carayannis, Evangelos Grigoroudis and Yorgos Goletsis(2016)[12] chose different indicators when comparing and analyzing the innovation efficiency at the national level and the regional level, such as the total number of lifelong learning participants, the total R&D expenditure, high-tech exports, new product sales, patents and works, and the number of trademark registrations.

To sum up, scholars choose different indicators in different industries when determining innovation efficiency indicators. However, for the vast majority of enterprises, R&D personnel and R&D funds are used as input index variables, and the sales revenue of R&D products and the number of books obtained are used as output index variables. Therefore, based on the research results at home and abroad, this paper adopts the following evaluation system, as shown in Table 1.

Table 1 Evaluation System of Technological Innovation Efficiency of Big Data Enterprises

Primary index	Secondary index	Three-level index	Three-level index measurement
Innovation investment	Manpower index	research staff	Proportion of R&D personnel (%)
	Funding index	R&D investment	Ratio of R&D investment to operating income (%)
		fixed assets	Net fixed assets of LN (yuan)
Innovative output	scientific and technical payoffs	Number of patents	Invention patent (piece)
	economic benefits	profit	Total profit of LN (yuan)
		invisible assets	Increase of intangible assets in LN (RMB)
envionment variables	macroscopic view	Marketization function	Growth rate of sales revenue (%)
	microcosmic	Scale	Total assets of LN (yuan)
		Ownership concentration	Proportion of the largest shareholder (%)

2.4 Measurement of Influencing Factors of Innovation Efficiency

In the daily production and operation process, there are many factors that will affect the innovation efficiency of big data enterprises. Therefore, in order to deeply analyze the factors and degrees that affect innovation efficiency, this section takes the comprehensive technical efficiency, pure technical efficiency and scale efficiency calculated by DEA model as dependent variables, and the social capital of the top management team as independent variables, and establishes Tobit regression analysis to explore the influence degree of the social capital of the top management team on innovation efficiency. The definition of innovation efficiency and its influencing factors are shown in Table 2.

Table 2 Definition of innovation efficiency and influencing factors variables

Statistical variable	Variable name	Variable symbol	Definition
Explained variable	Comprehensive technical efficiency	CRS	According to DEA innovation efficiency measurement
	Pure technical efficiency	VRS	According to DEA innovation efficiency measurement
	Scale efficiency	SCALE	According to DEA innovation efficiency measurement
Explanatory variable	Academic capital	Edu_c	Number of senior management team members who have served in universities, scientific research institutions, technical intermediaries, etc.
	Government capital	Gov_c	Number of senior management team members who have served in government units (including prefecture-level governments, people's congresses

		and CPPCC)
Guild capital	Gui_c	Number of senior management team members in trade associations
Number of senior management teams	TMT_n	Total number of senior management team members

To a certain extent, the social capital of TMT reflects the ability of enterprises to acquire and integrate resources embedded in social networks, which is one of the important factors affecting organizational innovation performance. According to different research purposes, different scholars have different measurement methods for TMT social capital. Based on the annual reports of listed companies, this paper manually sorts out the social capital of the top management team of 55 companies and enterprises. This thesis divides the social capital of the top management team into the following three categories: academic capital (Edu-c), government capital (Gov-c) and guild capital (Gui-c). Academic capital is mainly measured by the number of senior management teams who have served in universities and scientific research institutions; Political capital is measured by the number of senior management team members who have served in government units or have served in government units; Guild capital is measured by the number of senior management team members holding leadership positions in trade associations. This paper intends to explore the "black box" in the innovation process of big data enterprises and make an overall evaluation of the innovation efficiency of China's big data industry.

Using Tobit regression model to establish the relationship model between innovation efficiency of big data listed companies and influencing factors of social capital of senior management team, as shown in formulas (1), (2) and (3):

$$CRS_{i,t}=\beta+\beta_1 Edu\_c_{i,t}+\beta_2 Gov\_c_{i,t}+\beta_3 Gui\_c_{i,t}+\mu_{i,t} \tag{1}$$

$$VRS_{i,t}=\beta+\beta_1 Edu\_c_{i,t}+\beta_2 Gov\_c_{i,t}+\beta_3 Gui\_c_{i,t}+\mu_{i,t} \tag{2}$$

$$SCALE_{i,t}=\beta+\beta_1 Edu\_c_{i,t}+\beta_2 Gov\_c_{i,t}+\beta_3 Gui\_c_{i,t}+\mu_{i,t} \tag{3}$$

In the above formulas, β represents a regression constant; βi represents the regression coefficient; μ represents random error.

III. DATA ANALYSIS RESULTS

3.1 DESCRIPTIVE STATISTICS

According to the established index system, the data of relevant indicators are collected. Then the input indicators, output indicators and environmental variables of the selected 55 big data listed companies from 2016 to 2020 are statistically analyzed. The specific analysis results of each year are shown in Table 3-5.

Table 3 Descriptive statistics of input indicators

Index	Year	Obs	Mean	Min	Max	Std. Dev.
Proportion of R&D personnel (%)	2016	55	21.70	1.77	63.63	17.01
	2017	55	23.81	0.85	11.66	17.13
	2018	55	24.06	1.63	70.74	17.32
	2019	55	25.94	0.37	73.61	18.13
	2020	55	27.76	0.19	75.66	19.12
Ratio of R&D investment to operating income (%)	2016	55	5.67	0.00	27.73	5.57
	2017	55	6.27	0.00	26.24	6.57
	2018	55	6.43	0.01	31.83	6.44

	2019	55	6.54	0.02	27.29	5.95
	2020	55	6.52	0.05	27.54	5.70
Net fixed assets of LN (yuan)	2016	55	20.11	17.27	26.64	1.67
	2017	55	20.26	17.16	26.62	1.71
	2018	55	20.37	16.98	26.56	1.70
	2019	55	20.44	16.86	26.47	1.70

Descriptive statistics of input indicators. From 2016 to 2020, the average number of R&D personnel increased from 21.70 to 27.76, indicating that the number of R&D personnel in the personnel structure of big data listed companies is on the rise, but there is obvious imbalance among enterprises.

Descriptive statistical analysis of output indicators. There are significant differences in the number of invention patents of different companies, but the profitability of big data listed companies is better. The external environment has different impacts on the profitability of various enterprises, and the innovation ability of enterprises is increasing.

Table 4 Descriptive statistics of output indicators

Index	Year	Obs	Mean	Min	Max	Std. Dev.
Invention patent (piece)	2016	55	9.95	0.00	279.00	44.91
	2017	55	15.38	0.00	354.00	57.13
	2018	55	9.42	0.00	255.00	39.30
	2019	55	5.20	0.00	209.00	28.16
	2020	55	2.86	0.00	70.00	10.24
Total profit of LN (yuan)	2016	55	18.00	-21.28	22.42	7.56
	2017	55	19.59	-16.97	21.85	1.11
	2018	55	17.42	-22.03	23.21	9.41
	2019	55	16.67	-22.50	23.37	10.65
	2020	55	13.21	-21.63	23.49	14.36
Increase of intangible assets in LN (RMB)	2016	55	18.75	11.07	24.00	1.86
	2017	55	18.97	15.83	23.98	1.57
	2018	55	19.12	15.93	23.98	1.54
	2019	55	19.24	16.26	23.97	1.52
	2020	55	19.29	16.15	23.94	1.51

Descriptive statistical analysis of environmental variables. According to Table 3, the average growth rate of sales revenue from 2016 to 2020 showed a gradual decreasing trend, the average and standard deviation of total assets showed little difference, and the average and standard deviation of equity concentration showed a slow decreasing trend, which indicated that big data listed companies were generally stable.

Table 5 Descriptive statistics of environmental variables

Index	Year	Obs	Mean	Min	Max	Std. Dev.
Growth rate of sales revenue (%)	2016	55	0.60	-0.77	11.18	2.07
	2017	55	0.18	-0.85	2.06	0.45
	2018	55	0.13	-0.28	1.30	0.27
	2019	55	0.11	-0.43	0.95	0.26
	2020	55	-0.11	-0.95	0.51	0.25

Total assets of LN (yuan)	2016	55	22.66	20.46	27.15	1.16
	2017	55	22.81	20.61	27.08	1.11
	2018	55	22.94	20.96	27.02	1.06
	2019	55	23.02	21.08	27.06	1.03
	2020	55	23.10	21.20	27.09	1.05
Proportion of the largest shareholder (%)	2016	55	32.22	6.35	62.74	15.34
	2017	55	30.05	6.35	59.79	14.12
	2018	55	29.57	5.85	59.79	13.88
	2019	55	28.88	5.85	59.79	13.30
	2020	55	28.21	5.43	56.56	13.08

3.2 DEA INNOVATION EFFICIENCY ANALYSIS

This paper uses stata16.0 to measure the innovation efficiency of the panel data of various big data listed companies in China from 2016 to 2020. In this paper, the average value of DEA data envelopment analysis measures from 2016 to 2020 is counted by year, and the details of each year are shown in Table 6.

Table 6 Overall Analysis of Innovation Efficiency of Big Data Listed Companies

DEA efficiency	Efficiency eigenvalue	2016	2017	2018	2019	2020	Average value
Comprehensive technical efficiency	average value	0.935	0.923	0.916	0.912	0.931	0.923
（Effch）	TE=1 number of enterprises	11	12	10	eight	eight	-
Pure technical efficiency	average value	0.948	0.936	0.934	0.930	0.946	0.939
（Pech）	TE=1 number of enterprises	14	16	14	12	14	-
Scale efficiency	average value	0.986	0.985	0.981	0.981	0.984	0.983
（Sech）	TE=1 number of enterprises	17	17	10	eight	14	-
Scale income	Increasing the number of enterprises	39	27	18	20	nine	21
	Decreasing the number of enterprises	five	16	27	28	38	23
	Constant number of enterprises	11	12	10	seven	eight	10

From the time dimension, there are 9 companies with the best technical efficiency and scale every year, but the number of big data listed companies in the best state every year is small. Most companies are still in the stage of increasing returns to scale. Therefore, resources should be appropriately increased, so as to improve the innovation efficiency and make more innovative breakthroughs.

From the dimension of efficiency decomposition, the average scale efficiency is 0.983, the average pure technical efficiency is 0.939 and the average comprehensive technical efficiency is 0.923. This shows that the average comprehensive technical efficiency still has room for improvement; The technical level of big data listed companies is relatively backward, and the resources invested have not been effectively utilized; The scientific research management system of big data listed companies needs to be improved to effectively and efficiently control the whole innovation process, and at the same time optimize the input and output, so that the technical level can be in line with the head enterprises.

From the perspective of scale income, the average number of big data listed companies with constant scale income is 10. The average number of big data listed companies in increasing returns to scale is not much different from that of declining big data listed companies, which clearly shows that the overall efficiency level of big data listed companies in China is low. The structure of resource input and output resources of big data listed companies needs to be further optimized.

3.3 MALMQUIST INDEX ANALYSIS

The static efficiency of big data listed companies is calculated by DEA-BCC model, and then the dynamic changes of technological innovation efficiency of big data listed companies from 2016 to 2020 are further calculated by Malmquist index (Table 7- Table 8). From the results of dynamic changes, it can be concluded that from 2016 to 2020, the overall technological innovation efficiency of listed companies with big data concept on the main board is relatively high; Second, according to the relationship between the efficiency and the fluctuation of the efficiency value, we can draw the conclusion that the efficiency is not stable enough.

Table 7 Average Malmquist Index of 55 Big Data Listed Companies from 2016 to 2020

Code	Total factor productivity	Technical efficiency	Efficiency of technological progress	Pure technical efficiency	Scale efficiency
603888	0.993	0.994	0.999	0.993	1.001
603881	0.996	1.002	0.994	0.991	1.011
603869	0.984	0.988	0.995	0.990	0.998
603660	0.980	0.983	1.001	0.986	0.997
603636	1.009	1.016	0.993	1.000	1.016
603559	0.998	0.992	1.005	1.000	0.992
603508	0.986	0.979	1.010	0.988	0.989
603038	1.000	1.012	0.989	1.010	1.002
603019	0.935	0.994	0.946	0.988	1.005
603000	0.990	0.996	0.994	0.998	0.998
601789	0.974	0.993	0.981	0.996	0.997
601519	1.014	1.023	0.993	1.009	1.013
600996	0.995	0.997	0.998	1.034	0.969
600936	0.984	1.001	0.985	1.001	1.000
600894	0.988	0.997	0.992	1.003	0.994
600850	1.001	1.000	1.001	1.000	1.000
600845	0.995	0.998	0.998	1.005	0.993
600831	0.998	0.981	1.024	1.000	0.981
600804	0.999	1.009	0.991	1.008	1.001
600797	0.974	0.979	0.996	0.979	1.000
600756	1.004	0.998	1.008	0.998	1.000
600728	1.017	1.002	1.015	1.001	1.001
600633	0.985	0.986	0.999	0.986	1.000
600602	1.008	1.009	0.998	1.012	0.998
600590	1.001	0.997	1.005	0.992	1.005
600588	1.006	1.010	0.997	1.008	1.002
600522	0.994	0.999	0.996	1.015	0.986
600498	0.997	0.992	1.006	0.993	0.998
600487	0.988	0.996	0.993	0.998	1.000
600410	0.997	1.006	0.991	1.002	1.003
600271	0.917	1.001	0.909	1.000	1.001

600251	0.878	1.005	0.874	1.000	1.005
600166	0.844	1.014	0.830	1.000	1.014
600100	0.977	0.998	0.980	0.996	1.001
600070	0.988	0.996	0.992	0.996	0.999
600050	0.985	0.990	0.997	1.000	0.990
600037	0.998	1.001	0.997	1.003	0.998
000997	0.981	0.982	1.001	0.981	1.001
000977	0.999	1.001	0.998	1.010	0.992
000948	1.018	1.020	0.998	1.001	1.020
000938	0.997	1.000	0.997	1.000	1.000
000925	0.989	0.998	0.992	0.998	1.000
000889	0.993	0.996	0.996	0.997	1.000
000861	1.389	1.000	1.389	1.000	1.000
000851	0.983	0.996	0.987	1.000	0.996
000815	1.038	1.037	1.001	1.003	1.032
000711	1.001	1.010	0.992	1.010	1.000
000555	1.009	1.014	0.996	1.011	1.003
000547	1.006	1.008	1.000	1.005	1.003
000530	0.996	1.005	0.992	1.004	1.000
000158	0.991	0.998	0.993	1.007	0.992
000156	0.997	1.005	0.993	1.008	0.997
000070	0.995	0.999	0.997	0.999	1.000
000066	0.951	0.991	0.964	0.991	1.000
000032	0.987	1.000	0.987	1.000	1.000
Average value	0.994	1.000	0.995	1.000	1.000

Table 8 Malmquist index and its decomposition in average year from 2016 to 2020

Age	Total factor productivity	Technical efficiency	Efficiency of technological progress	Pure technical efficiency	Scale efficiency
2016-2017	1.015	0.989	1.026	0.988	1.001
2017-2018	0.969	0.993	0.977	0.997	0.996
2018-2019	1.010	0.995	1.016	0.997	0.998
2019-2020	0.985	1.023	0.962	1.019	1.005
Average value	0.995	1.000	0.995	1.000	1.000

3.4 Tobit Regression Results and Analysis of Influencing Factors of Innovation Efficiency

Based on the above analysis, the Tobit regression model is established with stata16.0, and the results are as follows:

According to Table 9, the social capital of TMT has a positive impact on the comprehensive technical efficiency of big data listed companies. The results of regression analysis of influencing factors of comprehensive efficiency show that academic capital, government capital, guild capital and the number of senior management teams have statistical significance on the total sample model, in which academic capital, government capital and guild capital have positive effects, while the total number of senior management teams has negative effects. Among them, guild capital has a more significant impact on the improvement of comprehensive technological innovation efficiency in big data listed companies.

Table 9 Influence of social capital of TMT on comprehensive technological innovation efficiency

Variable	Coefficient of regression	Standard Deviation	Z statistics	P value
Edu_c	0.001	0.002	1.320	0.013**
Gov_c	0.006	0.002	2.930	0.004***
Gui_c	0.013	0.003	4.400	0.001***
TMT_n	-0.003	0.001	-3.880	0.000***

Note: * * and * * * mean significant at the level of 5% and 1% respectively.

Table 10 shows that academic capital, government capital and guild capital will have a significant impact on pure technical efficiency. Compared with the impact on comprehensive technical efficiency, academic capital has a more significant impact on the efficiency of pure technological innovation. Among the social capital of top management team, academic capital, government capital and guild capital all have positive effects, and guild capital plays a more significant role in improving innovation efficiency.

Table 10 Influence of social capital of TMT on pure technical efficiency

Variable	Coefficient of regression	Standard Deviation	Z statistics	P value
Edu_c	0.001	0.002	1.500	0.621***
Gov_c	0.007	0.002	3.350	0.000***
Gui_c	0.013	0.003	4.190	0.000***
TMT_n	-0.003	0.001	-4.000	0.000***

Note: * * and * * * mean significant at the level of 5% and 1% respectively.

As can be seen from Table 11, among the three measures of social capital of TMT, only guild capital has a significant positive impact on the scale efficiency of big data listed companies.

Table 11 Influence of social capital of TMT on scale efficiency

Variable	Coefficient of regression	Standard Deviation	Z statistics	P value
Edu_c	0.001	0.001	1.050	0.296
Gov_c	0.001	0.001	1.280	0.203
Gui_c	0.002	0.001	1.830	0.068*
TMT_n	0.000	0.000	-0.57	0.568

Note: *, * * and * * respectively mean significant at the level of 10%, 5% and 1%.

IV. DISCUSSION

4.1 CONCLUSION

Based on 55 listed companies of big data concept stocks, this thesis further analyzed the innovation efficiency and its influencing factors of big data industry in China at this stage by using data envelopment analysis, Malmquist index and Tobit regression model. The conclusions are as follows.

Firstly, the big data industry plays a vital role in promoting China's innovation capability. In addition, China's big data companies are still in the initial stage of development, and there is still a certain gap compared with western developed countries. Therefore, in order to further improve the development level of China's digital economy, we should focus on the development of big data companies and promote the significant improvement of technological innovation efficiency of big data companies.

Secondly, through DEA data envelopment analysis, this study finds that most companies in China will be in a state of increasing or decreasing returns to scale due to unreasonable resource allocation. Therefore, it is urgent to further optimize the level of company resource allocation to improve innovation efficiency. Through Malmquist index analysis, it can be concluded that the technological investment of big data companies in recent years needs to be further improved, and breakthrough technological innovation will empower the innovation efficiency of big data companies. Using Tobit regression model, this paper finds that guild capital has a significant impact on comprehensive technological innovation efficiency, scale innovation efficiency and pure technological innovation efficiency, and academic capital and government capital are also the key factors for enterprises to improve their technological innovation efficiency. The improvement of technological innovation efficiency can enhance the comprehensive technological efficiency of big data listed companies to a greater extent.

4.2 PRACTICAL IMPLICATIONS

In order to improve the innovation efficiency of big data listed companies in China, this study will put forward suggestions from three angles: government, big data industry and big data listed companies themselves, in order to bring about more industrial innovation.

First of all, as a government, we should strive to create a social environment for innovation and change, and promote the strategy of strengthening the country through talents. Secondly, in the big data industry, it is necessary to improve the management level of the whole industry and seek pilot technology and talents. Finally, for big data companies, it is necessary to improve their innovation efficiency by improving the rational allocation of resources, accelerating the introduction of innovative talents, and upgrading and maintaining the social capital of senior management teams.

4.3 FUTURE RESEARCH

In the future research work, this study will, through its own practical experience, hope to make a more profound and thorough interpretation of the results of the analysis at this stage, explore the main reasons for the overall low innovation efficiency of large-scale big data listed companies in China, and analyze the decisive factors that affect the innovation efficiency of large-scale big data listed companies in China from two aspects: the overall environment that affects the development of big data industry, external factors and factors that affect big data companies themselves and internal factors. At the same time, the amount of data selected in this paper is still limited, and I hope to select more samples and data for empirical analysis in the future work, so as to draw more powerful conclusions.

REFERENCES

[1] Kim K Y, Lim C Y, Kim E J. A new approach to the space－time analysis of big data: application to subway traffic data in Seoul[J]. *Journal of Big Data*, 2018,25(1):132-140.

[2] Ilyasova N, Kupriyanov A, Paringer R and Krish D. Particular use of BIG DATA in medical diagnostic tasks[J]. *Pattern Recognition and Image Analysis*, 2018, 28: 114-121.

[3] Huang A.The Era of Artificial Intelligence and Big Data Provides Knowledge Services for the Publishing Industry in China[J].*Publishing Research Quarterly*,2019,35(1):164-171.

[4] Amankwah-Amoah J. Safety or no safety in numbers？Governments,bigdata and public policy formulation[J].*Industrial Management& Data Systems*,2015,115(9);1596-1603.

[5] Edwards J S, Taborda E R. Using knowledge management to give context to analytics and big data and reduce strategic risk[J]. *Procedia Computer Science*, 2016, 99: 36-49.

[6] Ferrier G D, Trivitt J S. Incorporating quality into the measurement of hospital efficiency: a double DEA approach[J]. *Journal of Productivity Analysis*, 2013, 40: 337-355.

[7] Doran J, McCARTHY N, O’connor M. The importance of internal knowledge generation and external knowledge sourcing for sme innovation and performance: evidence from Ireland[J]. *International Journal of Innovation Management*, 2019, 23(07): 1950069.

[8] Cruz-Cázares C, Bayona-Sáez C, García-Marco T. You can’t manage right what you can’t measure well: Technological innovation efficiency[J]. *Research Policy*, 2013, 42(6-7): 1239-1250.

[9] Kang K N, Park H. Influence of government R&D support and inter-firm collaborations on innovation in Korean biotechnology SMEs[J]. *Technovation*, 2012, 32(1): 68-78.

[10] Taghizadeh S K, Jayaraman K S and Ismail I. Innovation value chain as predictors for innovation strategy in Malaysian telecommunication industry[J]. *Problems and Perspectives in Management*, 2014, 12(4): 533-539.

[11] Nasierowski W, Arcelus F J. On the efficiency of national innovation systems[J]. *Socio-Economic Planning Sciences*, 2003, 37(3): 215-234.

[12] Carayannis E G, Grigoroudis E and Goletsis Y. A multilevel and multistage efficiency evaluation of innovation systems: A multiobjective DEA approach[J]. *Expert Systems with Applications*, 2016, 62: 63-80.