


# Research on the Development Level of Big Data Industry and Its Contribution to the Economy in China's Provinces

Sen Li<sup>1,2</sup>, Hongyuan Liang<sup>1</sup>, Yu Yan<sup>1</sup> and Haiying He<sup>1</sup>  
<sup>1</sup>Management School, Shenyang Jianzhu University, Shenyang, China  
<sup>2</sup>Institute of Applied Ecology, Chinese Academy of Sciences, Shenyang, China

**Keywords:** Big Data, Economic Growth, Heterogeneity.

**Abstract:** Based on the endogenous growth theory of economics, this paper puts forward that labor, capital and technology are the main driving forces to promote economic development. As the symbolic industry of the fourth technological revolution, big data has been integrated into every corner of the society and become the core technology element to promote regional economic development. This paper uses judgment matrix and other methods to calculate the development level of big data in each province. In order to study the contribution of big data industry to regional economy, this paper constructs a multiple regression model, taking the regional economic development level as the explained variable, the big data development level as the core explanatory variable, and labor and capital investment as the control variable. It is found that the big data industry has become an important force to promote regional economic development. At the same time, there is heterogeneity in the promotion efficiency of various regions. On this basis, it puts forward some suggestions on how to accelerate the development of big data industry and promote the coordinated development of China's provincial economy strategies and suggestions.


## 1 INTRODUCTION

Big data is a high-tech industry, and China will focus on its development as a strategic industry. Big data is the comprehensive application of modern technologies such as the Internet and cloud computing to classify and organize massive amounts of data, thereby providing solid data support for national management, social governance, and corporate management. Compared with European and American countries, China's big data industry started late, but after entering the 21st century, China's mutual big data industry has made great progress, especially in the fields of Internet shopping, mobile payment, e-commerce logistics, etc. Alibaba, JD, Tencent and other companies have become world-renowned companies in their respective fields, and their business scope covers most countries and regions in the world. The big data industry has become an important force in promoting the steady and healthy development of China's economy. (Wang 2021) believes that big data industry can not only promote the development of regional economy, but also

provide more scientific support for economic decision-making and help people find the best solution in complex social economy. Therefore, the development quality of big data industry has an important impression on the sustainable development of economy.

At the same time, we should note that there are obvious regional differences in the development level of China's economy. The development level of the southeast coastal areas is relatively high, and the development level of the inland areas is relatively low. The development of China's big data industry also has a similar situation. Therefore, from the perspective of the development of the big data industry, we analyze the heterogeneity of China's provincial economic growth momentum, and then propose countermeasures and suggestions to promote the coordinated development of various regions.

Big data is an emerging industry. At present, from the perspective of big data industry, it is rare to study the impact of big data industry on regional economic and industrial development. For example, (Zhang, 2020) proposed that big data industry is an important

 <https://orcid.org/0000-0002-4391-2582>

cornerstone of China's digital economy in the future, which can drive the transformation and upgrading of other industries. (Wu, et al., 2020) used the research method of big data to predict the growth of local economy. (Zhu, et al., 2020) proposed that big data industry can optimize the socialist market economy, and as a tool of macro-control, it can achieve a win-win situation between "government" and "market". At the same time, with the increasingly significant trend of economic globalization, the role of big data in international trade is becoming more and more prominent. (Zhang, 2021) proposed that information technology plays an important role in the operation of multinational enterprises. Big data provides important decision-making basis for enterprises in import and export trade. At the same time, big data can also help enterprises explore overseas markets and understand the different product needs of various countries.

The big data industry can not only become a new driving force for regional economic development, but also drive the transformation and upgrading of traditional entrepreneurship. (Wang, 2020) believes that based on big data analysis, traditional industries can more accurately find out the changes loved by consumers and the trend of market development, which can help them improve product competitiveness and enhance the profitability of enterprises.

It can be seen from the above literature that the current academic circles generally agree that big data industry is one of the key areas of future economic and industrial development, which can inject new power into the development of regional economy. At the same time, most of the existing studies are based on theory, and rarely use the method of data analysis to conduct empirical research on the contribution of big data industry to the economy. Therefore, based on the establishment and evaluation of the development level of big data industry in various regions of China, this paper will study the heterogeneity of China's economic development under the background of big data.

## 2 MODEL CONSTRUCTION AND DATA SOURCE

### 2.1 Model Construction

The empirical analysis of this paper is divided into two parts. The first part is to evaluate the development level of China's provincial big data industry. Build the index evaluation model of the development level of

big data industry, collect and sort out relevant data to evaluate the development level of big data industry in China's provinces and cities.

In order to improve the scientific evaluation of the development level of big data industry in various provinces and cities, this paper, based on the analytic hierarchy process, consults experts and scholars in the field of big data, re-determines the weight of the three dimensions, and then evaluates the development level of big data industry in various provinces and cities in China.

After the judgment matrix is constructed, the weight of each index is further determined. The formula is as follows:

$$W_i = \frac{1}{n} \frac{\sum_{j=1}^n a_{ij}}{\sum_{k=1}^n a_{ki}} \tag{1}$$

In the analytic hierarchy process, after obtaining the index weight, it is necessary to further test the consistency of the results. The calculation process is as follows:

Firstly, Calculate the consistency index C.I.,

$$C.I. = \frac{\lambda_{\max} - n}{n - 1} \tag{2}$$

Among them,  $\lambda_{\max}$  is the largest characteristic root of the matrix constructed in the previous section, and  $n$  is the number of indicators.

Secondly, look up the table to determine the corresponding average random consistency index R.I. (random index).

Table 1: Average random consistency index R.I.

Matrix order	1	2	3	4	5
R.I.	0	0	0.52	0.89	1.12

Finally, calculate the consistency ratio C.R. (consistency ratio), and compare the calculated consistency result with 0.1. If  $C.R. < 0.1$ , the result is considered to have passed the consistency test.

The second part is the heterogeneity analysis of China's provincial economic development, which constructs a multiple regression model, takes the development level of big data as an important influencing factor, and analyzes the differences in the driving factors of China's provincial and municipal economic development.

In this article, we use the level of development of big data as an indicator of technological progress. Therefore, we constructed a model of influencing factors of China's economic development.

$$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon \tag{3}$$

Among them,  $x_1$  is capital investment,  $x_2$  is human investment,  $x_3$  is the big data development index, and  $y$  is GDP.

### 2.2 Data Sources

All the data are from *China Statistical Yearbook 2019* and *Big data Blue Book: China big data development report No.3*. All indicators are cross-sectional data in 2018, and the analysis software is spss19.0.

## 3 DEVELOPMENT LEVEL OF BIG DATA INDUSTRY IN EACH PROVINCE

To evaluate the development level of big data industry, we need to collect relevant data of big data industry firstly. At present, as an emerging industry, the big data industry is still lack of relevant official statistics. Therefore, this paper get the data from *Big data Blue Book: China big data development report No.3*, which is compiled by the Key Laboratory of big data strategy. In this report, the development level of big data industry in each province is evaluated from three dimensions of government application(GM), commercial application(CE) and civil application (CI) from the perspective of big data application. However, in this report, the weights of the three ways of use are simply treated, the weights of the three are equal, and the sum of the three is taken as the comprehensive evaluation result, as shown in the table below.

Table 2: Big data application level index.

Region	GM	CE	CI	Region	GM	CE	CI
BEIJING	24.9	25.2	24.1	HAINAN	13.6	9.8	13.1
GUANGDONG	26.5	22.3	20.6	SHAANXI	12.9	11.2	11.9
ZHEJIANG	14.1	19.3	20.2	YUNNAN	11.1	9.4	15.5
SHANGHAI	22.1	15.5	15.7	HUNAN	12.3	12.4	11.1
GUIZHOU	28.9	7.7	16.2	NINGXIA	8.9	8.4	17.6
JIANGSHU	10.8	20.3	18.4	HUBEI	12.3	11.2	10.6
CHONGQING	20.7	11.7	14.5	QINGHAI	9.7	6.8	14.7
TIANJIN	19.7	11.6	15.5	SHANXI	12.6	5.7	11.0
SHANDONG	15.9	15.2	13.4	JIANGXI	10.8	9.4	8.7
HEBEI	18.6	10.6	13.6	JILIN	9.7	3.9	14.9
HENAN	22.3	9.0	11.5	GANSU	11.3	6.9	9.6
LIAONING	18.9	7.42	14.2	GUAGNXI	12.7	5.2	9.8
FUJIAN	11.3	12.5	14.7	HEILONGJIANG	9.5	3.9	12.8
ANHUI	14.1	14.5	9.4	XINJIANG	5.9	3.7	9.8
SICHUAN	13.4	12.3	12.2	XIZANG	1.5	7.4	1.7
NEIMENGGU	18.0	5.1	13.8				

As can be seen from table 2, there are great differences in the level of Internet application in various regions. In terms of government application, Guizhou, Guangdong and Beijing have the highest index; In terms of commercial application, Beijing, Guangdong and Jiangsu rank high in the index; In terms of XX application, Beijing, Guangdong and Zhejiang rank among the top three in the index. At the same time, the indexes of these provinces have exceeded 20, indicating a high level of development. The three Internet indexes of western inland provinces such as Tibet and Xinjiang did not exceed 10, ranking the last two of all provinces and cities.

After consulting experts in the field of big data, we obtained the judgment matrix, and the results are shown in the table below.

Table 3: Judgment matrix.

	GM	CE	CI
GM	1.00	0.20	0.33
CE	5.00	1.00	2.00
CI	3.00	0.50	1.00

After calculation, the weights of government applications, commercial applications and civil applications are  $(0.1096, 0.5813, 0.3092)^T$ , which shows that experts generally believe that the application of big data in the commercial field is the most important. At the same time, the value of C.R. is 0.0036, so the previous calculation result is correct.

Through the above calculation, we get the weight of the big data development level indicator, and then we can determine the big data development index of each province and city. The calculation results are as follows:

Table 4: Big data comprehensive development index.

Region	Score	Region	Score	Region	Score
BEIJING	24.80	LIAONING	10.78	QINGHAI	9.56
GUANGDONG	22.23	FUJIAN	13.07	SHANXI	8.11
ZHEJIANG	19.01	ANHUI	12.86	JIANGXI	9.35
SHANGHAI	16.25	SICHUAN	12.41	JILIN	7.95
GUIZHOU	12.68	NEIMENGGU	9.17	GANSU	8.25
JIANGSHU	18.67	HAINAN	11.26	GUAGNXI	7.42
CHONGQING	13.53	SHAANXI	11.60	HEILONGJIANG	7.22
TIANJIN	13.68	YUNNAN	11.46	XINJIANG	5.85
SHANDONG	14.71	HUNAN	12.00	XIZANG	4.99
HEBEI	12.41	NINGXIA	11.29		
HENAN	11.22	HUBEI	11.39		

As can be seen from the Table 4, there is a large gap in the development level of big data industry

among China's provinces. Beijing, Guangdong, Zhejiang, Jiangsu and other economically developed regions have higher big data index values and rank high, while Tibet, Xinjiang, Heilongjiang and other economically underdeveloped regions have lower values and rank last. Therefore, it can be preliminarily judged that there is a correlation between the development of big data industry and the level of regional economic development, which needs to be further tested by regression analysis.

#### 4 ANALYSIS ON THE CONTRIBUTION OF BIG DATA TO THE ECONOMY

There is a certain gap in the level of economic development in various regions of China. In order to facilitate research, according to the classification method of the China Bureau of statistics, China's provinces and cities are divided into four parts: the eastern region, the central region, the western region and the northeast region. In traditional economic theory, capital investment and human investment are important forces to promote economic growth and industrial development, and technological progress is another important factor.

Meanwhile, since China's provinces and cities are divided into four regions, in the regression analysis, China, the eastern region, the central region, the western region and the northeast region are taken as the research objects to study the heterogeneity of the contribution of big data to the economic growth of various regions in China.

##### 4.1 Descriptive Statistical Analysis

Before performing regression analysis, first perform statistical descriptive analysis on relevant data. According to the previous division of provinces and cities in China, the average value of the four main indicators in each region is calculated. The calculation results are shown in Table 5.

Table 5: Descriptive statistical analysis.

	GDP (Ten billion)	Human Investment (Million)	Capital Investment(Ten billion)	Big Data Index
East	481.00	20.22	619.12	16.61
Central	321.10	12.50	361.52	10.82
West	153.59	6.25	191.46	9.85
Northeast	189.17	7.32	228.62	8.65

It can be seen from Table 5 that there are obvious gaps in the economic development levels of various regions in China. The eastern region has obvious advantages, its total economic output is even twice that of the western region or the northeast region. The central and northeastern regions rank second and third, and the western region ranks last. At the same time, human resources, resources and big data index also showed the same situation. The eastern and central regions ranked the top two, and the western and northeastern regions ranked the bottom two.

##### 4.2 Regression Analysis

After descriptive statistical analysis, further regression analysis of the heterogeneity of economic development in various regions of China from the perspective of the big data industry is carried out. As shown in Table 6, the goodness of fit and F value of the four regression models are at a high level, and the significance of each regression coefficient is good.

Table 6: Regression analysis results.

	East	Central	West	Northeast
C	9807.43*** (2.863)	-9814.26** (3.453)	-2415.46 (-0.725)	- 6579.711 (-0.925)
$\beta_1$	5.045** (2.324)	4.136*** (2.912)	12.313*** (3.660)	26.337*** (4.234)
$\beta_2$	0.612*** (7.925)	0.550** (2.312)	0.588*** (5.897)	1.804*** (3.453)
$\beta_3$	589.01*** (2.497)	171.338*** (4.321)	120.65*** (6.234)	167.87*** (3.654)
Adjusted R <sup>2</sup>	0.998	0.697	0.929	0.926
F	423.41	57.7481	382.699	134.34

Note: the standard deviation in brackets, \*\*\*, \*\*, \* indicate the significance level of 1%, 5%, and 10% respectively.

From Table 6, it can be seen that the driving forces of economic growth in various regions of China are obviously heterogeneous. On the whole, the big data industry has played a key role in the economic development of various regions. The influencing factor of the level of big data development has been tested in the four models, which can show that the development of the big data industry has a positive impact on the growth of the regional economy, and the result passed the test. At the same time, in the eastern region, the big data industry contributes the most to the development of the regional economy, the coefficient is 589.01; while the western region has the

least contribution to the regional economy, the coefficient is 120.65. This is also consistent with the current distribution of the development level of the big data industry in various regions in China. Therefore, in order to accelerate the economic development level of China's central and western regions and northeast regions and promote the coordinated development of China's economy, the development of the big data industry in these regions should be accelerated.

## 5 CONCLUSIONS AND RECOMMENDATIONS

In this paper, Judgment matrix, multiple linear regression analysis are adopted to study the development level of big data industry and its contribution to regional economy. The main conclusions can be summarized as follows:

(1) There are great differences in the development level of big data industry among provinces and cities. The comprehensive development level of big data industry is high in economically developed provinces such as Beijing, Guangdong and Zhejiang. At the same time, it also ranks high in subdivided fields such as government application, commercial application and civil application. The development of big data industry in inland provinces such as Xinjiang, Tibet and Heilongjiang is relatively poor and has great room for improvement.

(2) The development of big data industry has a positive impact on regional economic growth. At the same time, among the four regions, the big data industry contributes the most to the economic development of the eastern region; The contribution to the western region is the smallest, and the contribution to the central region and the northeast region ranks second and third, which is also in line with the current distribution of the development level of big data industry in various regions of China.

Based on the above empirical analysis results, in order to accelerate the development of big data industry and promote the high-quality economic development of provinces and cities, the following countermeasures and suggestions are put forward.

(1) The big data industry should be included in the key development plan of the 14th five year plan. At present, Chinese governments at all levels are preparing the 14th five year plan. As a medium and long-term development plan, big data industry should be regarded as a key industry for governments at all levels to develop in the next 5-10 years. Support the

big data industry in terms of funds and policies. Especially in the central and western regions, the local government should fully understand the importance of big data industry, and strive to achieve leapfrog economic development through big data industry.

(2) All regions should take big data enterprises as key investment targets. Big data related enterprises have high added value, green environmental protection and other outstanding advantages, which meet the requirements of national high-quality development. All regions should attract and cultivate big data related enterprises in combination with the current situation of economic and industrial development. Actively learn from Guizhou, Inner Mongolia and other inland provinces the experience of big data industry investment, and promote the coordinated development of China's big data industry.

(3) Further strengthen the big data talent training. Big data industry is a technology intensive industry, and its high-quality development is inseparable from excellent talents in the field of big data. At present, some universities in eastern provinces have opened big data related majors, which have trained a large number of talents for the society. Therefore, the central and western provinces should also add big data related majors in time according to the needs of the market to meet the development needs of the local big data industry.

## ACKNOWLEDGMENT

This work was financially supported by 1. Liaoning Federation of Social Sciences—Study on the coupling of coordinated development of real estate and urban economy in Liaoning Province (NO.2022lslwtk-049) 2. Department of Education of Liaoning Province—Research on coupling development of strategic emerging industries and traditional industries in Liaoning Province (NO.lnqn202031).

## REFERENCES

- AA Anoushiravani, J Patton, Z Sayeed, MM El-Othmani, KJ Saleh (2016). Big Data, Big Research: Implementing Population Health-Based Research Models and Integrating Care to Reduce Cost and Improve Outcomes. *Orthopedic Clinics of North America*, vol. 4, pp:717-724.
- Anthony G. Picciano (2012). *The Evolution of Big Data and Learning Analytics in American Higher Education*.

- J. Journal of Asynchronous Learning Networks, vol. 3, pp:9-20.
- D Blazquez, & Domenech, J. (2018). Big data sources and methods for social and economic analyses. *Technological Forecasting and Social Change*, 130, 99-113.
- DA Alter (2015). Merits and Pitfalls of Using Observational "Big Data" to Inform Our Understanding of Socioeconomic Outcome Disparities. *J. Journal of the American College of Cardiology*, vol. 17, pp:1898-1900.
- Jiantian ZHU, Yanxia HE (2020). Big Data: Reshape the "New Planned Economy" or Improve the Socialist Market Economy? *J. West forum*, vol. 5, pp:13-21.
- Li, D. F., & Liu, P. D. (2020). Big data and intelligent decision methods in economy, innovation and sustainable development. *Technological and Economic Development of Economy*, 26(5), 970-973.
- Manyika J, Chui M, Brown B, et al (2011). *Big data: The next frontier for innovation, competition, and productivity*. USA: Mckinsey Global Institute.
- Ruihui WU, Xiangzhen ZHOU (2020). Research on short-term prediction of local economy based on big data. *J. Journal of University of Science and Technology Liaoning*, vol. 4, pp:304-308.
- Wang, S. (2021). An interview with shouyang wang: research frontier of big data-driven economic and financial forecasting. *Journal of Information Technology and Data Management*.
- Wang, X. M. (2020). Research on the application of big data statistics in the field of economic management. *Sci-tech Innovation and Productivity*.
- Xiuchun Zhang (2020). Big data is an important cornerstone of China's digital economy in the future. *J. Information China*, vol. 4, pp:56-59.
- Xue, B, Xiao, X, & Li, J. (2020). Identification method and empirical study of urban industrial spatial relationship based on poi big data: a case of shenyang city, china. *Geography and Sustainability*, 1(2).
- Zhang, Y. (2021). Research of the impact of big data on enterprise import and export based on economic globalization. *E3S Web of Conferences*, 253(2), 02001.
- Zheng, S. (2021). Influence of Big Data Development on Regional Rural Economic Growth.