ISSN 2616-5902 Vol. 6, Issue 8: 184-191, DOI: 10.25236/AJBM.2024.060827

The Impact of Data Element Aggregation on Green Economy Development: Evidence from China

Yaqi Wang

School of Business Administration, Northeastern University, Shenyang, 110000, China

Abstract: *Green economic development was put forward in 2015, and the 2030 Agenda for Sustainable Development was agreed by the United Nations. For today's world economic and social transformation, two major trends include digital economy and green economy. A quasi-natural experiment was used in this study, and panel data from 30 provinces, autonomous regions and municipalities from 2011 to 2022 was the basis of the study. The DID model can be constructed, and the development of green economy can be studied under the influence of data. In order to promote the development of green economy, three approaches can be used by data element aggregation: optimizing the ecological environment, enhancing technological levels, and facilitating enterprise digital transformation. Heterogeneity analysis reveals that in provinces with lower levels of green finance development, the influence of data element aggregation on green economy development is significantly enhanced. Therefore, it is crucial to focus on data element aggregation to promote green economic development.*

Keywords: *Data Elements, Green Economy, Big Data Comprehensive Pilot Area*

1. Introduction

Due to the continuous development of digital technology, China's economy's high-quality development can be promoted, and the productivity of data elements is released because of intelligent transformation and digitalization. The Action Plan to promote the development of Big Data has been put forward by The State Council, and the pilot construction of the pilot area of national Big data comprehensive could be started by Guizhou Province first, and "carrying out regional pilot" was clearly proposed, and the Beijing-Tianjin-Hebei, Pearl River Delta, Shanghai, Henan and other regions were expanded in October 2016. The boom of the digital economy is caused by the big data integrated pilot zone policy, and the experience can be provided for the development of big data across the country.

In the practice of various Big Data Comprehensive Pilot Areas, significant effects could be observed in industrial digitization, digital industrialization, and digital governance. The low-carbon and sustainable economic characteristics of data elements can promote the green transformation and upgrading of the three major industries. The aggregation of data elements driven by the Poilt Area policies has significant theoretical significance and research value for the green economy's development .

2. Literature Review

China is in a critical stage of transformation and development, with data being a new production factor in the digital economy era, serving as important support for constructing a new development pattern. The digital economy possesses environmentally friendly characteristics, reflected in improving urban air quality, driving regional innovation ecosystem niche fitness[1]. It is particularly important that the regional low-carbon development model's change is promoted by the zone of national Big Data comprehensive pilot [2].

How is the green economy affected by the aggregation of various factors discussed in the literature. The aggregation of factors such as technology and human capital has significant positive linear effects on green economic development[3], while the aggregation of financial factors exhibits an inverted U-shaped non-linear effect[4]. Additionally, the spatial spillover effects released by capital factor aggregation can empower green economic development[5]. Furthermore, existing research has focused on the greening effects of industrial synergistic aggregation, including pollution reduction, resource optimization, and enhancement of green total factor productivity[5], providing references for exploring the enabling role of data element aggregation in the development of green economy . Specifically, data

ISSN 2616-5902 Vol. 6, Issue 8: 184-191, DOI: 10.25236/AJBM.2024.060827

elements can release the benefits of green economy development through ecological governance channels that integrate information from different levels of ecological governance[2].

Regional economic growth requires not only the aggregation of advantageous elements but also innovation-driven ecological economic development. From the current development point of view, how the green economy's development affects the aggregation of data elements will not be studied. However, data element aggregation can enable high-quality development by optimizing labor markets[6], enhancing enterprise production efficiency[7], and improving the ecological environment[1]. It can be concluded that quasi-natural experiments can be built by the national big data comprehensive pilot zone, the mechanism of green economy development and how the influence effect is affected by the aggregation of data elements can be studied, and the research on green economy empowerment can be expanded.

3. Research hypothesis and theoretical analysis

The "14th Five-Year Plan for Digital Economy Development" issued by the State Council of China clearly states: "The digital economy is the main economic form following agricultural and industrial economies, promoting the unity of fairness and efficiency through data resources as key elements, forming a new economic form. "

3.1 Direct Impact of Data Element Aggregation on Green Economy Development

A sustainable development model, green economic development can enable efficient development and low carbon, which can reduce ecological damage[8]. Data element aggregation can directly impact green economic development from different dimensions. First, it promotes low-carbon development. Data element aggregation can aid in the construction of low-carbon cities by reducing information asymmetry, improving information transmission efficiency, and increasing information transparency. Second, it promotes efficient development. Data elements can enhance energy utilization efficiency, guide the comprehensive utilization of resources, and optimize energy consumption structures, driving high-efficiency economic and social development. Overall, data element aggregation has a unique synergistic effect on the elements of the sustainable development framework, driving low-carbon, efficient, and sustainable economic development.

3.2 Indirect Impact of Data Element Aggregation on Green Economy Development

The Big Data Comprehensive Pilot Area promotes the aggregation of data-intensive industries represented by the big data industry, optimizing the ecological environment and driving green development. First, data elements are clean and sustainable, enhancing total factor productivity and resource allocation efficiency, achieving maximum output with minimal pollution emissions. Second, big data industry aggregation can fully integrate and utilize data elements, improving ecological environment information networks, and organically uniting data resources with ecological environment construction. Additionally, deep exploration of data element applications can innovate ecological regulation, significantly improving urban air quality, enterprise pollution control levels, and optimizing energy resource allocation[9].

The Big Data Comprehensive Pilot Area, through establishing big data centers and introducing leading enterprise innovation centers, explores the innovative application of big data. First, data element aggregation, through the collection, mining, and analysis of data elements, can discover information value, achieve an informed appreciation, and promote digital technology progress[10]. Second, data elements can enable traditional production factors, optimize technical routes, and accelerate technological iteration[11]. Therefore, when the development of green economy is promoted, digital technology plays a driving role.

The Big Data Comprehensive Pilot Area adopts coordinated development measures which have a positive effect on accelerating industrial digitization and data element marketization. First, data elements are crucial factors in digital economy development. Data elements link various segments of the digital economy's industrial and innovation chains, producing a multiplier effect through aggregation[12]. Second, data element aggregation positively impacts enterprise innovation development, facilitating enterprise innovation development by alleviating labor resource mismatches, increasing R&D investment levels, and improving human capital levels[13]. Third, in the changing digital economy of regional enterprises, the positive policy-driven role can be played by the big data integrated pilot zone policie,

ISSN 2616-5902 Vol. 6, Issue 8: 184-191, DOI: 10.25236/AJBM.2024.060827

further integrating big data with physical industries[14].

In summary, data element aggregation promotes green economic development through channels such as optimizing the ecological environment, advancing digital technology, and driving industrial digital transformation.

4. The Design of Research

4.1 The Construction of Model

When explaining how the development of green economy plays a role in the aggregation of data elements, the China Big Data Comprehensive Experimental Area is the research basis of this study. The Comprehensive Poilt Area of National Big Data Development Report (2018 edition) is the research basis of this study, and the experimental area includes Tianjin, Hebei, Guangdong, Shanghai, Henan, Chongqing, Shenyang and other regions, with the remaining provinces as the control group. Since the traditional difference-in-differences (DID) method cannot study situations where the experimental group policies are implemented at different time points, this study constructs a time-varying DID model as follows:

$$GEEV_{it} = \alpha_0 + \alpha_1 DID_{it} + \beta_2 Z_{it} + e_i + y_t + \varepsilon_{it} \tag{1}$$

where $GEEV_{it}$ represents the green economy of region;*i* and *t* are region and time dimentions, respectively; the dummy variable DID_{it} as the core explanatory variable represents $treat_{it} \times time_{it}$; if province *i* was approved to establish a big data conprehensive pilot zone in 2015 or 2016, then the variable $time_{it}$ takes the value of 1; otherwise, it takes the value of 0; Z_{it} donates control variables; e_i and y_i represents individual fixed effects and time fixed effects, restectively; ε_{it} is the random disturbance term.

Additionally, to examine the mechanism through which data element aggregation impacts greeneconomy development, this study constructs a mediation effect model referencing Jiang Ting (2022)[15].

$$Mediator_{it} = \alpha_2 + \alpha_3 DID_{it} + \beta_3 Z_{it} + e_i + y_t + \varepsilon_{it} \tag{2}$$

where $Mediator_{it}$ serves as the mediation variable, representing the ecological environment, technological level, and enterprise digitalization level, Equation (1) is consistent with other variables.

4.2 Variable Selection

4.2.1 Explained Variable

The dependent variable is green economy. Existing literature measures green economy levels primarily through two approaches: an indicator system based on green GDP and efficiency measurement. Currently, the construction of indicator systems to measure green economy development is influenced by scholars' subjectivity, which may lead to inaccurate measurement results. Following the principles of systematicity, feasibility, and scientific rigor, this study comprehensively considers resource inputs and environmental costs to construct green economy efficiency, referencing the research of He Weida[16]. The green economy measurement indicators are shown in Table 1.

Table 1: Green Economy Efficiency Measurement Indicator

Indicator Type	Variable	Indicator Definition
Green Energy Efficiency	Regional Output	Regional Gross Domestic Product (billion yuan)
	Regional Energy Consumption	The Consumption of Energy (10, 000 tons of standard coal)
Green Environmental Efficiency	Wastewater Discharge Efficiency	Ratio of regional wastewater discharge to regional output (10, 000 tons)
	Gas Emission Efficiency	Ratio of regional output to regional industrial emissions of sulphur dioxide and particulate matter (tons)
	Solid Waste Emission Efficiency	Ratio of regional output to the regional production industrial solid waste emissions (tons)

ISSN 2616-5902 Vol. 6, Issue 8: 184-191, DOI: 10.25236/AJBM.2024.060827

4.2.2 Core Variable

The core explanatory variable is data element aggregation. Following the methodology of Liu Chuanming (2023)[17] , The proxy variable of data element aggregation is the big data integrated pilot area policy, represented as a dummy variable. For Guizhou Province, the variable is set to 0 before 2015 and to 1 from 2015 onwards; for the other nine provinces, the variable is set to 0 before 2016 and to 1 from 2016 onwards. 1 and 0 are assigned to the experimental group and the control group.

4.2.3 Mediator Variables

(1)Optimize the ecological environment: This study constructs an indicator system to measure ecological, environmental levels from two aspects: pollutant emissions and pollution control. For pollutant emissions, Data such as per capita industrial wastewater discharge, per capita industrial dust discharge and per capita sulfur dioxide discharge can be extracted. In the process of pollution control, indicators such as environmental treatment input,data such as the comprehensive utilization rate of industrial solid waste and the harmless treatment rate of domestic waste can be used.

(2)Raise the level of science and technology: R&D expenditure intensity is listed as an important indicator in monitoring the progress of China's innovative nation-building process. In order to measure the level of technology, the R&D expenditure’s intensity can be used in this paper, and the R&D’s ratio capital stock to regional GDP is represented. The R&D expenditure price index is calculated by weighting the Consumer Price index (CPI) and the fixed asset price index, with weights of 45% and 55% respectively. The base year is 2001, the actual internal R&D expenditure can be eliminated, and 15% is the depreciation rate of the R&D capital stock.

(3)Improve the digital level of enterprises: This study constructs an indicator system to measure enterprise digitalization levels. The number of computers used per 100 employees and the number of websites per 100 companies are included in three indicators built into the system, and entropy is the main way to measure how digital a company is.

4.2.4 Control Variables

Considering that green economy development is influenced by other factors, this study selects the following variables as control variables, based on the characteristics of digital economy :

(1)Development Level of Infrastructure : The level of regional infrastructure development positively impacts China's total factor productivity, which may affect green economic development.

(2)Regional Economic Development: Differences in regional economic development can influence green economy efficiency.

(3)Rationalization of Industrial Structure: Rationalizing the industrial structure can improve regional ecological governance and promote high-quality economic growth.

(4)Government Intervention: Market resource allocation can be promoted through appropriate government intervention, and positive effects will be produced on the development of green economy.

4.3 The Defination and The Descriptive Statistic of Variable

The definitions and descriptive statistics of variables are shown in Table 2.

Table 2: The Defination and the Descriptive Statistic of Variable

Variable Type	Variable Symbol	Variable Defination	N	Mean	Std.	Min	Max
Explained variable	Green Economy (*GEEV*)	Green economy efficiency, constructed using the factor decomposition method referenced from He Weida (2022)	360	5.875	6.591	0.734	62.225
Explanatory variable	Data Element Aggregation (*DID*)	Proxy variable represented by the Big Data Comprehensive Pilot Area policy	360	0.194	0.396	0	1
Control variable	Infrastructure Development (*PT*)	Measured with the per capita number of public transportation vehicles in the area	360	12.777	2.957	7.05	26.55
	Regional Economic Development (*INGDP*)	Regional per capita GDP (lnpgdp), log-transformed	360	5.872	3.066	1.602	19.031
	Industrial Structure	Measured using the Theil index	360	0.150	0.937	0.007	0.451

ISSN 2616-5902 Vol. 6, Issue 8: 184-191, DOI: 10.25236/AJBM.2024.060827

	(*IFM*)						
	Government Intervention (*LECDP*)	The proportion of local fiscal expenditure to local GDP can be measured	360	0.259	0.112	0.105	0.758
Mediator	Ecological Environment (*NEV*)	Constructed based on the research of Yue Yujun (2024)	360	0.749	0.132	0.425	0.989
	Technological Level (*RDGDPF*)	Intensity of R&D expenditure	360	0.017	0.012	0.004	0.082
	Enterprise Digitalization (*CDA*)	Constructed based on the research methods of Wang Jun (2021) and Zhao Tao (2020)	360	0.120	0.096	0.013	0.960

5. Explain the empirical results

5.1 Baseline regression

The baseline regression results of data element aggregation on green economic development are shown in Table 3, columns (1) and (2). In column (1), time-fixed effects and region-fixed effects can be controlled. The results of regression indicate that data element aggregation has a stimulating effect on regional green economy development. In column (2), four control variables are added: infrastructure development, the development level of regional economy, the structure of industry, and government intervention level. The results show that the estimated coefficients for infrastructure development and industrial structure are significantly negative, while the estimated coefficients for regional economic development level and government intervention level are positive.

5.2 Mechanism Test

This study explores the influence mechanism of data element aggregation on the green economy's development through the mediating effects of ecological environment, technological level, and enterprise digitalization level. Table 3 (3), (4) and (5) show the research results. Firstly, the effect of data element aggregation on the ecological environment is 0.024 and significant. Secondly, the impact coefficient of data element aggregation on technological level is 0.001 and signs. Lastly, the effect of data element aggregation on enterprise digitalization level is 0.038 and signs. In summary, data element aggregation can improve the ecological environment, achieve information value-added, enhance technological levels, and boost enterprise digitalization levels, thereby promoting green economic development.

Table 3: Baseline Regression and Mechanism Test and Heterogeneity Test

Variable	*GEEV*	*GEEV*	Ecological Environment	Technology level	Enterprise digitalization level	Higher green finance development	Lower green finance development
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
DID	4.851*** (4. 86)	2.447*** (5.26)	0.024*** (2.637)	0.001*** (2.554)	0.038*** (3.828)	0.282 (0.40)	2.137*** (8.54)
PT		-0.438*** (-3.54)	0.005*** (2.645)	-0.000*** (-2.481)	-0.000 (-0.040)		
IFM		-12.238* (-2.30)	-0.026*** (-4.890)	0.002*** (6.630)	0.029*** (5.101)		
INGDP		3.236*** (6.62)	0.138 (1.613)	-0.012*** (-2.572)	0.041 (0.602)		
RECDP		27.977*** (6.48)	-0.092 (-0.919)	0.005 (0.917)	0.280*** (3.647)		
_cons	5.157*** (21.52)	-13.189*** (-4.11)	0.835*** (20.277)	0.012*** (4.258)	-0.135** (-2.117)	-23.027***(-3.88)	-6.600*** (-6.91)
Year	YES	YES	YES	YES	YES	YES	YES
Region	YES	YES	YES	YES	YES	YES	YES
N	360	360	360	360	360	180	180
R^2	0.804	0.915	0.976	0.976	0.691	0.943	0.967

Standard errors are in parentheses. The signifcance level at 10%, 5%, and 1% is denoted, respectively, by *, **, and ***.

5.3 Heterogeneity Test

To eliminate the disparities caused by different levels of green finance between provinces in the Big Data Comprehensive Experimental Zone policy, this study measures green finance levels using green

ISSN 2616-5902 Vol. 6, Issue 8: 184-191, DOI: 10.25236/AJBM.2024.060827

credit and divides the sample into high and low green finance level groups. It could been seen from the test results of heterogeneity in columns (6) and (7) of Table 3, for economically poor provinces, the impact of big data integration point policies is positive, If the level of green finance in the province is high, there is no significant positive impact.

5.4 Parallel Trend Test

The prerequisite for DID is parallel trend test, which conducted using Jacobson's event study method. The results in Figure 1 show that prior to the policy implementation. The policy coefficient of the big data integrated experimental area enables this non-significance to be demonstrated, and the significant difference does not exist between the control group and the experimental group. After the policy was implemented, the coefficients show an upward trend, indicating a significant policy effect.

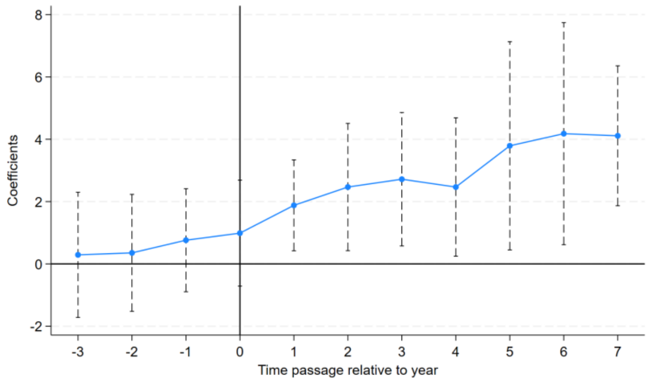

Figure 1: Parallel Trend Test

5.5 Placebo Test

The placebo test can be randomly assigned to the treatment group.The results in Figure 2 show that the policy estimation coefficient of the treatment group's big data comprehensive experimental area is close to 0, which is consistent with the normal distribution, and the robustness of the experiment can be explained.

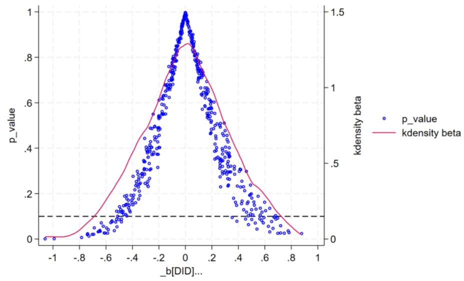

Figure 2: Placebo Test

5.6 Analysis of Experimental Results

5.6.1 Endogeneity Test

Considering that the model might encounter endogeneity issues due to reverse causality, measurement errors, and other problems, Endogeneity test can be carried out by the use of instrumental variable (IV) method in this study.Following the methodology of Huang Qunhui (2019)[18], The instrumental variables of this study can be reflected by the interaction terms of the number of fixed-line telephones per 100 people in each province and the amount of national Internet investment, thus enabling the two-stage estimation to be carried out. In Table 4, the F-value of Cragg-Donald Wald in columns (1) and (2) is greater than the 10% critical value of Stock-Yogo, which indicates that it passes the weak instrument variable test. In addition, at the 1% significance level, the null hypothesis can be rejected by the K-Paap rk LM value, proving that the identifiability test is passed. The positive and significant impact of data element aggregation on green economy development remains, demonstrating the stability of the results.

5.6.2 Robustness Test

To test the robustness of the impact of data element aggregation on green economic development, this study employs two methods based on existing literature:

Replacing the Explained Variable: In column (3) of Table 4, the data element aggregation's regression coefficient and green economy development is positively correlated, and the robustness of the results can be confirmed. The concentration level of data elements can be measured by the computer service's proportion and the industry employees of software and information transmission at the end of the period, so that the robustness test can be carried out.

Single DID Test: Although the Big Data Comprehensive Pilot Area policy was piloted in 2015, its peak was in 2016. This study reselects 2016 as the policy starting year to test the effectiveness of the policy evaluation. It can be seen from column (4) that at the 1% level, the positive significance can be reflected by the regression coefficient, and the stability of the results can be demonstrated.

Table 4: Endogeneity and Robustness Test

Variable	The regression of first-stage	The regression of second-stage	Altrenative explained variable	Single DID
	(1)	(2)	(3)	(4)
DID		17.074*** (3.322)	2.804*** (4.36)	1.363*** (3.67)
iv	0.000*** (2. 90e-06)			
Cragg-Donald Wald F statistic		31.373		
K-Paap rk LM statistic		38.888		
Year	YES	YES	YES	YES
Region	YES	YES	YES	YES
N	360	360	360	360
R^2		0.692	0.932	0.924

6. Conclusion and Policy Recommendation

Panel data for 30 provinces from 2011 to 2022 were the data source for the study, this study constructs a multi-timepoint DID model to measure data element aggregation using policy of the Big Data Comprehensive Pilot Area.The study employs baseline regression, the test of parallel trend, the test of placebo, the test of robustness, mechanism test, and heterogeneity analysis to explore the data's impact element aggregation on green economy development. The findings indicate that data element aggregation significantly promotes green economy development through optimizing the ecological environment, enhancing technological levels, and facilitating enterprise digital transformation. The effect is more pronounced in provinces with lower economic development, lower green finance development, or higher technological advancement.

Policy recommendations include promoting data element aggregation to optimize the ecological environment, increasing investment to facilitate enterprise digital transformation, encouraging technological innovation, and implementing differentiated policies tailored to local conditions to maximize the benefits of data element aggregation and achieve the high-quality economy's development.

References

[1] He Wei, Dong Ying, Lü Jingwen. The Impact of Digital Transformation on the Niche Fitness of Regional Innovation Ecosystem [J]. Science & Technology Progress and Policy, [2024-06-07]: 1-10.
[2] Zhang Xiufan, Fan Decheng. Research on the Development of Digital Economy Empowering my country's Low-Carbon Economic Transformation--Based on the Analysis of the National Big Data Comprehensive Experimental Areas [J]. Science & Technology Progress and Policy, 2023, 40(19):118-128.
[3] Marrucci L, Daddi T, Iraldo F. The contribution of green human resource management to the circular economy and performance of environmental certified organisations[J]. Journal of Cleaner Production, 2021, 319: 128859.
[4] Deng D , Meng H , You M A ,et al.Effect of extracellular polymeric substances on Dolichospermum aggregation during temperature rise[J].Journal of Oceanology and Limnology, 2023, 41(6):2208-2218.

ISSN 2616-5902 Vol. 6, Issue 8: 184-191, DOI: 10.25236/AJBM.2024.060827

[5] Ma Xing, Xie Hua. Fiscal Decentralization, Industrial Synergy and Agglomeration, and UrbanGreen Development [J]. Journal of Southwest Minzu University(Humanities and Social Sciences Edition), 2023, 44(09):83-93.

[6] Xu L .Renewable resource of aggregation-induced emission materials: From photophysical mechanisms to biomedical applications[J].Coordination chemistry reviews, 2024, 506(May):1.1-1.20.

[7] Brynjolfsson E, Hitt L M, Kim H H. Strength in numbers: How does data-driven decisionmaking affect firm performance?[J]. Available at SSRN 1819486, 2011.

[8] Allouche J, Middleton C, Gyawali D. Technical veil, hidden politics: Interrogating the power linkages behind the nexus [J]. Water Alternatives, 2015, 8(1): 610-626.

[9] Xue Y, Tang C, Wu H, et al. The emerging driving force of energy consumption in China: Does digital economy development matter?[J]. Energy Policy, 2022, 165: 112997.

[10] Carriere-Swallow M Y, Haksar M V. The economics and implications of data: An integrated perspective [M]. International Monetary Fund, 2019.

[11] LIANG Lin, JIN Guangmin. Research on the path of digital economy to empower the resilience of China's industrial chain[J]. Qilu journal. 2023(5):129-138.

[12] OUYANG Rihui. Data as a factor of production promotes the deep integration of the digital economy and the real economy: Theoretical logic and analysis framework[J]. Economic Review Journal. 2024(2):67-78.

[13] El-Kassar A N, Singh S K. Green innovation and organizational performance: The influence of big data and the moderating role of management commitment and HR practices[J]. Technological forecasting and social change, 2019, 144: 483-498.

[14] Shi Yutang, Wang Xiaodan. Can the establishment of a Big Data Comprehensive Pilot Area drive thedigital transformation of enterprises?--Empirical research based on quasi-natural experiments [J/OL]. Studies in Science of Science: 1-18[2024-06-22].

[15] JIANG Ting. Mediating Effects and Moderating Effects in Causal Inference[J]. China Industrial Economics, 2022, (05):100-120.

[16] HE Wei-da, WEN Jia-long, et al. Research on the lmpact of Digital Economy Development on China's Green Ecological Efficiency: Based on Two-way Fixed Effects Model[J]. On Economic Problems, 2022(01):1-8+30.

[17] Liu Chuanming, Chen Liang, Wei Xiaomin. Impact of Data Element Agglomeration on Scientific and Technological Innovation: A Quasi-natural Experiment Based on Big Data Comprehensive Pilot Areas[J]. Journal of Shanghai University of Finance and Economics, 2023, 25(05):107-121.

[18] Huang Qunhui, Yu Yongze, Zhang Songlin. Internet Development and Productivity Growth in Manufacturing Industry: Internal Mechanism and China Experiences[J]. China Industrial Economics, 2019(08):5-23.