Research on the Impact of Data Elementalization on Low-Carbon Economic Development

Yanlong Qian

Guizhou University of Finance and Economics, Guiyang 550025, Guizhou, China

Abstract: Under the “dual-carbon” goals background, data elementalization is regarded as a key force in promoting green economic transformation and achieving low-carbon development. In-depth research on its specific role holds significant theoretical and practical importance. Based on panel data from 30 provinces in China from 2012 to 2022, this paper constructs an econometric model to empirically examine the impact of data elementalization on the development level of the low-carbon economy. The study finds that data elementalization can significantly reduce carbon emission intensity and effectively enhance carbon total factor productivity. This positive promoting effect remains robust after endogeneity testing. Therefore, full attention should be given to the integration of data elements with the real economy to leverage their critical role in advancing green and low-carbon transformation.

***Keywords*: data elementalization, low-carbon economy, carbon emission intensity, carbon total factor productivity

1. Introduction

Global climate change has become a severe challenge, making the promotion of low-carbon economic development an international consensus. As one of the major economies, China, while sustaining economic growth, also faces significant pressure from high carbon emissions and urgently needs to propel a green transition of its economy. However, the current development model still heavily relies on traditional production factors, constrained by issues such as resource misallocation, efficiency bottlenecks, and insufficient innovation. In this context, the rapid development of the digital economy has established data as a key production factor. Its characteristics of being non-rivalrous, having low replication costs, and being shareable allow it to permeate all stages of production. By enhancing resource efficiency and optimizing energy systems, among other ways, it provides a new pathway for green, low-carbon transformation. Relevant policies also emphasize promoting the integrated application of data factors in related fields to leverage synergistic effects.

Existing research has predominantly focused on the overall level of the digital economy. There remains insufficient exploration into the intrinsic mechanisms of how data factors themselves influence the low-carbon economy, which limits the understanding and practical reference for the pathways of data factor capitalization. Therefore, systematically examining the impact of data factor capitalization on low-carbon economic development from the perspective of production factors holds significant theoretical and practical importance. This paper attempts to construct a measurement indicator for data factor capitalization and employs carbon emission intensity and carbon total factor productivity to comprehensively gauge the level of low-carbon economic development. It empirically tests the relationship between the two, aiming to provide references for related fields.

2. Literature Review

Current research on data factors is already quite extensive, primarily focusing on three aspects: First, the analysis of the concept and nature of data factors. Related studies posit that data is information that can be used to reduce prediction errors[1], and emphasize that, as a fundamental factor of production, it possesses technical-economic characteristics such as non-rivalry, low replication cost, non-excludability, externalities, and immediacy[2]. Second, the estimation of data factors. Existing research predominantly employs the cost method to estimate data factors. Xu et al. estimated the scale of China’s data factors using the cost method from the perspective of the information value chain[3]. Liu and Xiong estimated China’s data capital formation and stock through the cost method, value-added method, and storage scale method[4]. Third, the impact of data factors on economic development. At the micro level, studies incorporating data factors into production functions have found that their introduction in enterprise production can reduce uncertainty[5], enable the reorganization and reallocation of traditional factors[6], and empower ambidextrous innovation[7], thereby enhancing microeconomic efficiency and promoting economic development. At the macro level, existing research indicates that data factors primarily promote economic growth through multiplier effects. Furthermore, the output elasticity of data factors in the current stage far exceeds that of previous periods, making it one of the significant driving forces behind China's economic growth[8].

Existing literature mainly focuses on the impact of data factor capitalization on economic growth and efficiency, while relatively neglecting its role in the environmental dimension, particularly in low-carbon economic development. Therefore, research on how data factor capitalization affects the low-carbon economy needs further exploration.

3. Theoretical Analysis and Research Hypotheses

In the era of the digital economy, data has become a key factor of production, capable of linking innovation, activating capital, nurturing talent, and driving industrial upgrading and economic growth. With its unique advantages, the data factor exerts a profound impact on economic development at both macro and micro levels.

At the macro level, the role of the data factor is primarily reflected in two aspects. First, it promotes industrial upgrading and a shift in growth drivers. By integrating with industries, it generates digital dividends, fosters structural optimization, and transforms growth momentum. An increased level of data capitalization can accelerate the phasing out of heavily polluting industries and propel industries toward an environmentally friendly digital transformation[9]. Second, it enhances government governance efficacy. By empowering multi-stakeholder collaborative governance, it accelerates the transition from traditional government to digital government, improves the modernization level of governance, thereby strengthening the monitoring of environmental, resource, and economic operations, and supports low-carbon economic development.

At the micro level, data capitalization mainly functions in three ways. First, an improved level of data capitalization accelerates knowledge dissemination and the sharing of innovative data among enterprises, reducing R&D trial-and-error costs, thereby stimulating a large number of technological innovations and providing technical support for regional low-carbon development. Second, the combination of data with traditional factors drives enterprise digital transformation, mitigates factor misallocation, and enhances total factor productivity[10]. Simultaneously, by alleviating information asymmetry and reducing energy waste, it achieves the dual goals of carbon reduction and efficiency improvement. Third, an increased level of data capitalization heightens public environmental awareness. Leveraging the internet and big data platforms, the efficiency of pollution information dissemination improves, further strengthening public awareness of green consumption. Based on the above analysis, Hypothesis 1 is proposed:

H1: Data capitalization can enhance the level of low-carbon economic development in a region.

4. Empirical Model Specification and Variable Selection

4.1 Model Specification

To empirically examine the impact of data capitalization on the low-carbon economy, this paper constructs the following econometric model:

$$Y_{it} = \alpha + \beta DE_{it} + \gamma X_{it} + \theta_i + \delta_t + \varepsilon_{it} \tag{1}$$

Where I and t represent province and year, respectively; Y is the dependent variable, denoting carbon emission intensity and carbon total factor productivity; α is the model constant term; DEis the core explanatory variable, representing the level of data capitalization; β is the estimated coefficient for the core explanatory variable; X denotes the control variables; γ is the estimated coefficient for the control variables; θ irepresents province fixed effects; δ represents year fixed effects; and ε is the random error term.

4.2 Variable Selection

To comprehensively measure the level of low-carbon economic development, this paper selects carbon emission intensity (CI) and carbon total factor productivity (CTFP) as the dependent variables. Specifically, carbon emission intensity is expressed as the ratio of total carbon emissions to the region's actual GDP (using 2011 as the base year for conversion). Carbon total factor productivity is calculated using the Super-SBM model with undesirable outputs. The input factors include labor, fixed capital, and energy; the desirable output is the region's actual GDP; and the undesirable output is carbon emissions. More specifically, labor is measured by the year-end number of employed persons. Fixed capital is calculated using the perpetual inventory method, following Shan Haojie's study[11], with the year 2000 as the base period and a depreciation rate of 10.96%. Energy is measured by the total consumption of various energy sources. The carbon emission data in this paper are derived using the accounting method from Cong Jianhui et al.[12].

For the explanatory variable, data capitalization (DE), this paper follows the approach of Zhang Liao et al.[13] by using software business revenue as a proxy indicator for the level of data capitalization.

Furthermore, to control for endogeneity in the model, the following control variables are selected: Government Intervention (GI, fiscal expenditure / GDP), Population Density (PD, year-end resident population / administrative area), Energy Structure (EM, regional electricity consumption / national total electricity consumption), Financial Development Level (FDL, output value of the financial sector / GDP), and Social Consumption Level (SCL, total retail sales of consumer goods / GDP).

5. Empirical Results Analysis

5.1 Benchmark Regression Results

Table 1 reports the results of the benchmark regression. Columns (1) and (2) show that, whether control variables are included or not, the estimated coefficient of data capitalization on carbon emission intensity is significantly negative at the 5% level, indicating its ability to significantly reduce carbon emission intensity. Columns (3) and (4) demonstrate that, under both model specifications, the coefficient of data capitalization on carbon total factor productivity is significantly positive, suggesting it can significantly enhance carbon total factor productivity. In summary, data capitalization promotes low-carbon economic development at the provincial level through two dimensions—reducing carbon emission intensity and increasing carbon total factor productivity—thus validating Hypothesis H1.

Table 1. Baseline Regression

Variable name	(1)	(2)	(3)	(4)
	CI	CI	CTFP	CTFP
DE	-0.746^{**}	-0.705^{**}	0.078^{**}	0.100^{**}
	(-2.612)	(-2.708)	(2.177)	(2.535)
Control variable	No	Yes	No	Yes
P.Y. FE	Yes	Yes	Yes	Yes
N	330	330	330	330
R^2	0.919	0.925	0.896	0.912

Notes: *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively. Values in parentheses are t-statistics; the same applies hereinafter.

5.2 Endogeneity Test

To address potential endogeneity issues in the model, we further employ an instrumental variable approach for endogeneity testing. Drawing on existing research[14], the number of post offices per million people in 1984 is selected as the instrumental variable for the level of data capitalization. This historical variable is correlated with the current level of data capitalization while meeting the exogeneity requirement as a historical factor. In the actual estimation, its interaction term with the national information technology service revenue from the previous year is used as the final instrumental variable. Table 2 reports the results of the two-stage least squares (2SLS) estimation. Column (1) presents the first-stage regression results, showing that the instrumental variable (IV) is significantly correlated with the level of data capitalization, and the F-statistic exceeds 10, indicating no weak instrument problem. Columns (2) and (3) present the second-stage estimation results. They indicate that the direction of the impact of data capitalization on carbon emission intensity and carbon total factor productivity is consistent with the benchmark regression results. Specifically, it significantly reduces carbon emission intensity and enhances carbon total factor productivity, thereby confirming the robustness of the benchmark regression conclusions.

Table 2. Instrumental Variable Method

Variable name	(1)	(2)	(3)
	first-stage	second-stage	
	DE	CI	CTFP
DE_IV		-1.041^{***}	0.087^{***}
		(-3.803)	(2.672)
IV	0.000^{***}		
	(4.877)		
Control variable	Yes	Yes	Yes
P.Y. FE	Yes	Yes	Yes
Wald F	63.32		
N	330	330	330

6. Conclusions

Based on panel data from 30 Chinese provinces from 2012 to 2022, this study investigates the impact of data capitalization on low-carbon economic development. The results indicate that an increase in the level of data capitalization significantly reduces carbon emission intensity and enhances carbon total factor productivity (CTFP). The findings remain robust after including control variables and addressing endogeneity with an instrumental variable approach, confirming a causal relationship between data capitalization and low-carbon economic development.

This study offers the following insights for promoting low-carbon development: Policymakers should recognize the strategic value of data in the green transition. At the macro level, promoting the integration of data factors with the real economy is crucial for guiding industries toward digital and low-carbon transformation. Concurrently, governments should leverage data to enhance environmental governance through intelligent monitoring platforms. At the micro level, enterprises should be encouraged to utilize data-driven innovation and management optimization to reduce carbon footprints, while digital platforms can guide the public toward green lifestyles.Future efforts should focus on improving data market mechanisms to facilitate secure, orderly data flow and cross-regional collaboration, thereby fully harnessing the potential of data in achieving both high-quality economic growth and environmental protection.

References

[1] Farboodi M, Veldkamp L. Long-run growth of financial data technology[J]. American Economic Review, 2020, 110(8): 2485-2523.
[2] Cai Yuezhou, Ma Wenjun. Impact of data factors on high-quality development and constraints on data flow[J]. The Journal of Quantitative & Technical Economics, 2021, 38(03): 64-83.
[3] Xu Xiang, Tian Xiaoxuan, Li Keaobo, et al. Scale estimation and structural analysis of China's data factors: From the perspective of information value chain[J]. Contemporary Finance & Economics, 2024, (04): 3-16.
[4] Liu Taoxiong, Rong Ke, Zhang Yadi. Estimation of data capital and its contribution to China's economic growth: Based on the data value chain perspective[J]. Social Sciences in China, 2023, (10): 44-64+205.
[5] Farboodi M, Veldkamp L. A model of the data economy[R]. Cambridge, MA, USA: National Bureau of Economic Research, 2021.
[6] Acemoglu D, Restrepo P. Automation and new tasks: How technology displaces and reinstates labor[J]. Journal of economic perspectives, 2019, 33(2): 3-30.
[7] Li Li, Yong Hui, Wang Lei, et al. Data factor embedding, dual innovation and manufacturing transformation and upgrading[J]. Journal of Statistics and Information, 2024, 39(04): 31-45.
[8] Yang Yan, Wang Li, Liao Zujun. Market-oriented allocation of data factors and regional economic development: Based on the perspective of data trading platform[J]. Social Science Research, 2021, (06): 38-52.
[9] Chen Xiaojia, Xu Wei. Data factors, transportation infrastructure and industrial structure upgrading: Analysis based on quantitative spatial general equilibrium model[J]. Management World, 2024, 40(04): 78-98.
[10] Shi Dan, Sun Guanglin. Data factors and new quality productive forces: From the perspective of enterprise total factor productivity[J]. Economic Theory and Business Management, 2024, 44(04): 12-30.
[11] Shan Haojie. Reestimating the capital stock of China: 1952–2006[J]. The Journal of Quantitative & Technical Economics, 2008, 25(10): 17-31.
[12] Cong Jianhui, Liu Xuemin, Zhao Xueru. Boundary definition and measurement methods of urban carbon emission accounting[J]. China Population, Resources and Environment, 2014, 24(04): 19-26.
[13] Zhang Liao, Hu Zhongbo. Research on the impact of data factorization on common prosperity[J]. Soft Science, 2024, 38(11): 18-25+33.
[14] Huang Qunhui, Yu Yongze, Zhang Songlin. Internet development and manufacturing productivity improvement: Internal mechanism and China's experience[J]. China Industrial Economics, 2019, (08): 5-23.

Author Bio

Yanlong Qian (born August 2000), male, Wuhu, Anhui Province, Master's degree, research interests: digital economy.