Coupling Coordination between Corporate Digitalization and Green Efficiency under the Data Elements × Action Plan

Shiman Zhou *

School of Economics and Trade, Hunan University, Changsha 410082, China

* Corresponding Author Email: 1250688308@qq.com

Abstract. The release of the Data Elements × Three Year Action Plan (2024 to 2026) provides an important real world setting for studying the coordinated development of corporate digital transformation and green efficiency. Using panel data of Chinese A share listed manufacturing firms from 2014 to 2023, this paper constructs a composite corporate digitalization index covering five dimensions (artificial intelligence, big data, cloud computing, blockchain, and digital technology applications) based on annual report text analysis. Corporate green efficiency is measured by a super efficiency SBM model that incorporates undesirable outputs. On this basis, a coupling coordination degree model quantifies the dynamic synergy between the two subsystems, and a panel Tobit regression identifies key drivers of the coupling coordination degree. The empirical sample covers 2,683 listed firms and 21,328 firm year observations, with a complete pipeline of indicator standardization, entropy weighting, and coupling coordination grading. The main findings are as follows. First, the mean digitalization index rises from 0.142 to 0.387 over the sample period, with an annual growth rate of 10.5 percent. Second, mean green efficiency rises from 0.524 to 0.681, indicating steady improvement. Third, the mean coupling coordination degree rises from 0.412 to 0.586, moving the sample from the verge of disorder to the transition between barely coordinated and primary coordinated stages. Fourth, government R&D subsidy intensity, environmental regulation intensity, and firm human capital exhibit significantly positive effects on the coupling coordination degree. The paper provides a quantitative framework and policy implications for advancing the synergy between the Data Elements × Action Plan and green low carbon transformation of manufacturing.

Keywords: Data Elements × Action Plan, corporate digital transformation, green efficiency, coupling coordination degree, super efficiency SBM model.

1. Introduction

Data elements are widely regarded as the fifth fundamental factor of production after land, labor, capital, and technology, playing a critical multiplier role and serving as an innovation engine in the digital economy era. In January 2024, the National Data Administration jointly with 17 ministries released the Data Elements × Three Year Action Plan (2024 to 2026), which identifies 12 priority domains for releasing the value of data elements, among which Data Elements × Green Low Carbon serves as a key initiative for industrial energy saving, emission reduction, and green development [1]. Against the joint background of the dual carbon goals and the high quality development strategy, examining the coordinated evolution of corporate digitalization and green efficiency carries significant theoretical and practical implications for advancing manufacturing green transformation and building a digital ecological civilization.

Corporate digital transformation, as the core micro level response to the national digital economy strategy, has reshaped both production organization and value creation logic. Existing studies find that the digital economy raises regional total factor productivity through channels including industrial structure upgrading, mitigation of factor allocation distortions, and reduction of transaction costs[2]. At the firm level, scholars use the frequency of digital keywords in the Management Discussion and Analysis (MD&A) section of annual reports to construct firm level digitalization indices, and have empirically examined the effects of digital transformation on capital market performance, corporate division of labor, and total factor productivity [3-5]. These studies provide an important reference for the indicator system in this paper.

Green efficiency is a core indicator that jointly accounts for economic output and environmental costs, and its measurement methodology has undergone several iterations. The super efficiency Slacks Based Measure (SBM) model proposed by Tone introduces slack variables on top of conventional Data Envelopment Analysis (DEA) and allows further ranking of efficient decision making units. Once undesirable outputs are incorporated, the model captures the green performance of production processes more faithfully [6, 7]. Multiple studies apply this model to measure green total factor productivity at city and firm levels in China, and find that digital economy development significantly raises urban green total factor productivity, with particularly pronounced gains in resource based cities and central and western regions [8].

Recent research increasingly examines the interplay between digitalization and green development. On one hand, digitalization promotes green transformation by enabling green technological innovation, mitigating information asymmetry, and improving resource allocation efficiency [10-12]; on the other hand, government environmental regulation may crowd out corporate green innovation under certain conditions [13]. However, existing work largely adopts a one directional causal framework focused on the effect of digitalization on green performance, and rarely characterizes the coupling coordination between the two subsystems from a system synergy perspective. The coupling coordination degree model from physics provides a standardized method for multi system synergy analysis [9], yet its application at the corporate micro level remains limited.

To address the research gaps above, this paper proposes a firm level coupling coordination analysis framework for digitalization and green efficiency. The contributions are as follows. First, under the policy backdrop of the Data Elements × Action Plan, this paper applies the coupling coordination degree model to panel data of A share manufacturing listed firms to quantify the system synergy between corporate digitalization and green efficiency. Second, in terms of measurement, the paper combines annual report text analysis with a super efficiency SBM model with undesirable outputs to measure digital transformation and green efficiency respectively, enhancing the comparability and robustness of the two subsystem measures. Third, in terms of driver identification, a panel Tobit regression identifies key determinants of the coupling coordination degree and conducts heterogeneity analysis, providing empirical evidence for differentiated policy design [14]. Fourth, based on 21,328 firm year observations from 2014 to 2023, this paper systematically reveals the temporal evolution and regional / industry divergence patterns of the coupling coordination degree.

2. Methodology

To address key challenges including firm level digitalization measurement heterogeneity, nonlinearity in green efficiency calculation, and the difficulty of identifying multi system synergy, this paper constructs an empirical framework integrating text analysis, super efficiency DEA, coupling coordination modeling, and panel Tobit regression. The framework comprises two core stages: a data preparation and preprocessing stage, and a mathematical and statistical modeling stage. The former handles raw data collection, cleaning, and variable construction, while the latter performs composite evaluation of digitalization and green efficiency, computation of the coupling coordination degree, and identification of its drivers based on structured data.

2.1. Data Sources and Preprocessing

The empirical data are drawn from the CSMAR database, the Chinese Research Data Services platform (CNRDS), Wind Information, and the original text of annual reports of listed firms. The sample covers financial, patent, annual report text, and pollutant emission data for Chinese A share manufacturing listed firms from 2014 to 2023. After excluding ST firms, financial and insurance firms, and observations with missing key variables, the final sample contains 2,683 listed firms and 21,328 firm year observations. As shown in Table I, the general and special purpose equipment manufacturing industry and the computer, communications, and other electronic equipment manufacturing industry have the largest sample sizes, accounting for 13.2 percent and 11.8 percent of the total sample respectively. Chemical materials and products, pharmaceutical manufacturing, automobile manufacturing, and textile and apparel industries also have representative sample sizes, covering the major categories of Chinese manufacturing.

All raw datasets are matched by unified identifiers and preprocessed through the following steps.

1) Text data collection and cleaning: Annual report PDF files of listed firms are retrieved from the Shanghai and Shenzhen stock exchange websites using Python web crawlers. The Management Discussion and Analysis section is extracted with PyMuPDF as the analytical corpus, and garbled text and duplicate paragraphs are cleaned.

2) Keyword frequency statistics: A digital transformation terminology dictionary is built based on existing literature and national digital economy policy documents, covering 87 keywords across five dimensions (artificial intelligence, big data, cloud computing, blockchain, and digital technology applications). Word frequency is computed using the Python jieba segmentation tool. Negation contexts and references to entities other than the focal firm are excluded, producing annual keyword frequencies for each firm.

3) Input output variable construction and standardization: Input and output variables are adjusted to 2014 constant prices using industry specific price indices, and all continuous variables are winsorized at the 1 percent and 99 percent levels to mitigate the influence of outliers. All indicators are then nondimensionalized via the range method, and objective weights are determined by the entropy weighting method.

The preprocessed sample uses firm year observations as decision making units, yielding a panel of 21,328 rows. The composite digitalization index is built from 87 keywords across 5 dimensions, and green efficiency is built from 3 input indicators, 1 desirable output, and 2 undesirable outputs. This structured representation provides a consistent data foundation for subsequent model estimation.

2.2. Mathematical and Statistical Models

1) Indicator Standardization and Entropy Weighting

To unify dimensions and objectively reflect the variation of indicators across the sample, raw indicators are nondimensionalized via the range method. Let there be n sample firms and p evaluation indicators, with the raw value of the j th indicator for the i th sample denoted by x_{ij}. The standardization formulas for positive and negative indicators are given by Equation (1) and Equation (2) respectively.

$$x'_{ij} = \frac{x_{ij} - minx_j}{maxx_j - minx_j} \tag{1}$$

$$x'_{ij} = \frac{maxx_j - x_{ij}}{maxx_j - minx_j} \tag{2}$$

Where x'_{ij} denotes the standardized indicator value, max x_j and min x_j are the maximum and minimum values of indicator j across the sample, i indexes firms, and j indexes indicators. Based on the standardized indicator matrix, the entropy weighting method computes objective weights for each indicator. The normalized proportion of the j th indicator for the i th sample is defined in Equation (3).

$$p_{ij} = \frac{x'_{ij}}{\sum_{i=1}^{n} x'_{ij}} \tag{3}$$

The information entropy and difference coefficient of the j th indicator are computed by Equation (4) and Equation (5).

$$e_j = -\frac{1}{lnn}\sum_{i=1}^{n} p_{ij} lnp_{ij} \tag{4}$$

$$w_j = \frac{1 - e_j}{\sum_{k=1}^{p}(1 - e_k)} \tag{5}$$

Where e_j is the information entropy of the j th indicator, and w_j is the objective weight assigned by the entropy method. The composite development level U is computed by linear weighting, as shown in Equation (6).

$$U_i = \sum_{j=1}^{p} w_j x'_{ij} \tag{6}$$

U_i reflects the composite development level of the i th sample firm under the chosen indicator system. The procedure above is implemented separately for the digital transformation and green efficiency subsystems, producing the firm level composite digitalization index U_1 and the green efficiency composite index U_2.

2) Construction of the Corporate Digital Transformation Index

The Management Discussion and Analysis section of annual reports is segmented using the Python jieba tokenizer, and keyword frequencies are computed across the five dimensions (artificial intelligence, big data, cloud computing, blockchain, and digital technology applications). To accommodate variation in annual report length and the wide range of raw frequency values, the base digitalization indicator is defined as the natural logarithm of one plus the total keyword frequency, as in Equation (7).

$$Digital_{it} = ln(1 + \sum_{d=1}^{5} Freq_{d,it}) \tag{7}$$

Where $Freq_{d,it}$ denotes the sum of keyword frequencies for firm i in year t along dimension d. The base indicator, together with sub indices for underlying technology adoption and technology application practice, is fed into the entropy weighting system to produce the composite digitalization index U_1. U_1 lies in the interval [0, 1], with higher values indicating higher levels of digital transformation.

3) Super Efficiency SBM Model for Corporate Green Efficiency

Let there be n decision making units (firm year observations), each with m inputs, s_1 desirable outputs, and s_2 undesirable outputs. Inputs include labor (the firm's total employee count), capital (net fixed assets), and energy consumption (total energy converted to standard coal equivalent). The desirable output is operating revenue. Undesirable outputs are the firm's converted industrial sulfur dioxide emissions and carbon dioxide emissions. The super efficiency SBM model with undesirable outputs is given by Equation (8).

$$\rho = min \frac{1 + \frac{1}{m}\sum_{i=1}^{m} \frac{s_i^-}{x_{io}}}{1 - \frac{1}{s_1 + s_2}\left(\sum_{r=1}^{s_1} \frac{s_r^g}{y_{ro}^g} + \sum_{q=1}^{s_2} \frac{s_q^b}{y_{qo}^b}\right)} \tag{8}$$

Subject to the constraints in Equation (9).

$$x_{io} = \sum_{j=1, j\neq o}^{n} \lambda_j x_{ij} + s_i^-, y_{ro}^g = \sum_{j=1, j\neq o}^{n} \lambda_j y_{rj}^g - s_r^g, y_{qo}^b = \sum_{j=1, j\neq o}^{n} \lambda_j y_{qj}^b + s_q^b \tag{9}$$

Where ρ is the firm's green efficiency score computed by the model; x_{ij}, y_{rj}^g, and y_{qj}^b denote the i th input, the r th desirable output, and the q th undesirable output of the j th decision making unit; s_i^-, s_r^g, and s_q^b are slack variables for input redundancy, desirable output shortfall, and undesirable output excess; λ_j is the linear combination coefficient with λ_j greater than or equal to 0; all slack variables are also nonnegative. When ρ is greater than or equal to 1, the decision making unit is strongly efficient; when ρ is less than 1, it is weakly efficient or inefficient. The green efficiency composite index U_2 is set to min (ρ, 1.0) to ensure U_2 lies in [0, 1].

4) Coupling Coordination Degree Model

Drawing on the coupling coordination degree model of Liao [9], this paper builds the coupling coordination degree between the corporate digitalization subsystem (U_1) and the green efficiency subsystem (U_2). The coupling degree C between the two subsystems is defined in Equation (10).

$$C = 2\sqrt{\frac{U_1U_2}{(U_1+U_2)^2}} \tag{10}$$

The composite coordination index T is given by Equation (11).

$$T = \alpha U_1 + \beta U_2,\ \alpha + \beta = 1 \tag{11}$$

The coupling coordination degree D is the geometric mean of the coupling degree C and the composite coordination index T, as in Equation (12).

$$D = \sqrt{C \times T} \tag{12}$$

Where α and β are subsystem weights. Given the equal importance of digital transformation and green efficiency for corporate sustainable development, this paper sets α equal to β equal to 0.5. C reflects the strength of interaction between the two subsystems, T reflects the overall development level, and D reflects the degree of coordinated development. D values are partitioned into 10 levels: 0 to 0.1 extreme disorder, 0.1 to 0.2 severe disorder, 0.2 to 0.3 moderate disorder, 0.3 to 0.4 mild disorder, 0.4 to 0.5 verge of disorder, 0.5 to 0.6 barely coordinated, 0.6 to 0.7 primary coordinated, 0.7 to 0.8 intermediate coordinated, 0.8 to 0.9 well coordinated, 0.9 to 1.0 superiorly coordinated.

5) Panel Tobit Model for Driver Identification

Since the coupling coordination degree D lies in [0, 1] and is therefore a limited dependent variable, ordinary least squares may produce biased estimates. This paper adopts a panel Tobit model to identify drivers of the coupling coordination degree. Its latent variable formulation is given by Equation (13).

$$D_{it}^* = \beta_0 + \beta_1 RDS_{it} + \beta_2 ENV_{it} + \beta_3 HC_{it} + \sum_k \gamma_k Controls_{k,it} + \mu_i + \lambda_t + \varepsilon_{it} \tag{13}$$

The observed coupling coordination degree satisfies a piecewise censoring rule: when the latent variable lies in [0, 1] the observed value equals the latent value; otherwise it takes the boundary value, as in Equation (14).

$$D_{it} = max(0, min(D_{it}^*, 1)) \tag{14}$$

Where RDS is government R&D subsidy intensity (subsidy amount divided by operating revenue); ENV is environmental regulation intensity (regional environmental governance investment as a share of regional GDP); and HC is corporate human capital (the share of employees with a bachelor's degree or above). Control variables in Controls include firm size (log total assets), firm age, leverage ratio, and ownership attribute. μ_i denotes firm fixed effects, λ_t denotes year fixed effects, and ε_{it} is the random disturbance term. Cluster robust standard errors are used to mitigate potential within group correlation.

The above indicator standardization, entropy weighting, super efficiency SBM model, coupling coordination degree model, and panel Tobit regression form a consistent modeling pipeline: standardization and entropy weighting aggregate heterogeneous raw indicators into comparable composite indices; the super efficiency SBM model provides a rigorous operational research foundation for nonlinear green efficiency measurement; the coupling coordination degree model characterizes firm level dynamic synergy from a system perspective; and the panel Tobit regression further identifies the key drivers of that synergy.

3. Empirical Results

3.1. Descriptive Statistics

The empirical analysis is conducted on Stata 17 and Python 3.10, with the super efficiency SBM model implemented in MaxDEA Ultra. The full panel is randomly split into 80 percent training and 20 percent validation sets with a random seed of 42 for reproducibility; the validation set contains 4,266 firm year observations. Table I reports descriptive statistics for the main variables. The composite digitalization index U_1 has a mean of 0.214 and a standard deviation of 0.183, indicating substantial cross firm heterogeneity. Green efficiency U_2 has a mean of 0.612 and a standard deviation of 0.241. The coupling coordination degree D has a mean of 0.498, indicating a sample average in the transition from the verge of disorder toward barely coordinated.

Table I. Descriptive Statistics of Main Variables

Variable	Mean	Std Dev	Min	Max	N
Digitalization Index U_1	0.214	0.183	0.000	0.962	21328
Green Efficiency U_2	0.612	0.241	0.052	1.428	21328
Coupling Degree C	0.832	0.156	0.243	0.998	21328
Coordination Index T	0.413	0.165	0.034	1.105	21328
Coupling Coordination D	0.498	0.143	0.106	0.875	21328
R&D Subsidy RDS	0.018	0.022	0.000	0.213	21328
Env. Regulation ENV	0.013	0.007	0.002	0.041	21328
Human Capital HC	0.183	0.156	0.012	0.821	21328
Firm Size (lnAsset)	22.31	1.382	18.62	28.43	21328
Firm Age	17.42	6.183	1	53	21328
Leverage (Lev)	0.428	0.198	0.038	0.951	21328

Table I indicates that the overall digital transformation level of sample firms is relatively low but exhibits substantial heterogeneity, with a wide gap between the maximum and minimum values, reflecting uneven progress in digital transformation among Chinese listed firms. Mean green efficiency exceeds 0.6, indicating that most manufacturing listed firms exhibit relatively strong environmental performance, although improvement room remains. The mean coupling coordination degree sits near the 0.5 threshold, indicating significant potential for further synergistic improvement, which motivates the subsequent heterogeneity analysis by region, industry, and ownership.

3.2. Composite Levels of Digitalization and Green Efficiency

Figure 1 plots the temporal evolution of the mean digitalization index and mean green efficiency of sample firms from 2014 to 2023. The digitalization index rises steadily from 0.142 in 2014 to 0.387 in 2023, with an annual growth rate of 10.5 percent, accelerating noticeably after 2020. This pattern is consistent with policy drivers including the Data Elements × Action Plan and the digital transformation demand catalyzed by the COVID 19 pandemic. Mean green efficiency rises from 0.524 to 0.681, a slower but stable improvement, reflecting the continuous gains in corporate environmental performance under the dual carbon goals [8].

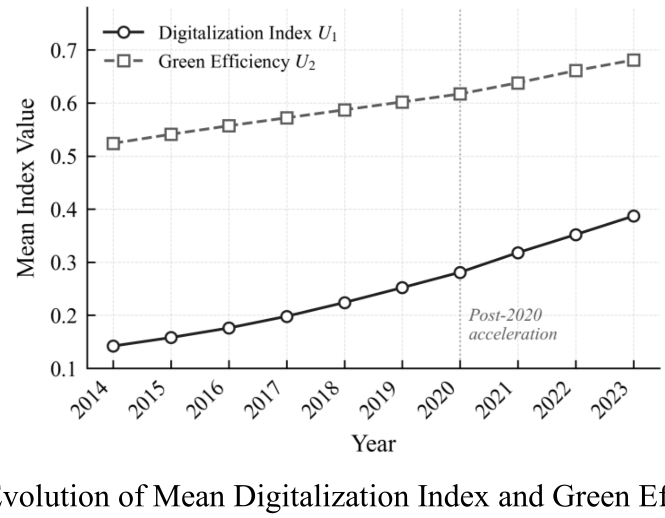

Figure 1. Temporal Evolution of Mean Digitalization Index and Green Efficiency, 2014 to 2023

Table II compares the mean digitalization index, green efficiency, and coupling coordination degree across major industry categories. High tech industries such as computer, communications, and other electronic equipment manufacturing exhibit a digitalization index significantly higher than that of traditional heavy industries such as steel and chemicals, while the gap in green efficiency between the two industry groups is relatively small. This suggests that traditional heavy industries have a certain foundation in energy use and pollution control but lag in digital transformation, and should be a focus of policy. In terms of the mean coupling coordination degree, the computer, communications, and other electronic equipment manufacturing industry reaches 0.582, already in transition between barely coordinated and primary coordinated, while ferrous metal smelting is only 0.452, still at the verge of disorder stage.

Table II. Mean Digitalization and Green Efficiency by Industry

Industry	Digitalization U_1	Green Eff U_2	Coupling Coord D	N
Computer, Comm. & Electronics	0.342	0.658	0.582	2517
General & Special Equipment	0.273	0.643	0.543	2814
Automobile Manufacturing	0.281	0.612	0.534	1462
Instruments & Meters	0.296	0.671	0.554	932
Pharmaceutical Manufacturing	0.226	0.694	0.515	1798
Chemical Materials & Products	0.184	0.594	0.471	2241
Ferrous Metal Smelting	0.158	0.586	0.452	643
Nonferrous Metal Smelting	0.171	0.598	0.463	871
Textile & Apparel	0.142	0.617	0.443	1387
Other Manufacturing	0.198	0.605	0.482	6663

3.3. Temporal and Regional Evolution of Coupling Coordination

Figure 2 traces the temporal evolution of the mean coupling coordination degree for the eastern, central, and western regions from 2014 to 2023. The eastern region rises from 0.451 in 2014 to 0.624 in 2023, leading the central and western regions throughout the sample period. The central region rises from 0.398 to 0.572 and the western region from 0.378 to 0.541. Regional gaps narrow but remain visible. In terms of grade transitions, the full sample mean evolves from the verge of disorder (0.412) to the transition between barely coordinated and primary coordinated (0.586), crossing two grades in total. This pattern is consistent with the spatial structure of Chinese digital infrastructure, where the east leads the west, and with the different paces of industrial structure transformation across regions.

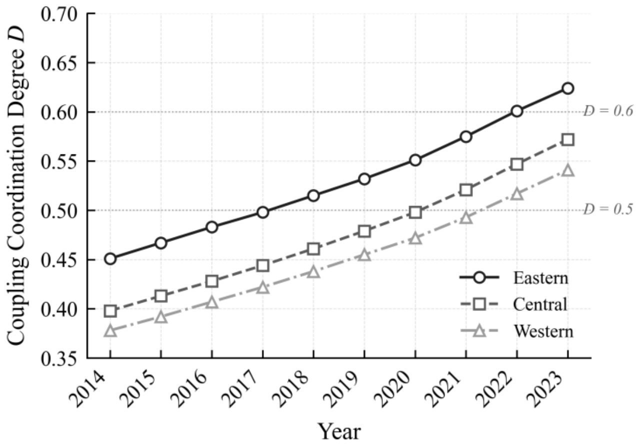

Figure 2. Temporal Evolution of Mean Coupling Coordination Degree across Eastern, Central, and Western Regions, 2014 to 2023

Further statistics on the coupling coordination grade distribution show that in 2014, 58.2 percent of sample observations were at the verge of disorder or below, while in 2023 this share dropped to 21.7 percent. Meanwhile, the share of observations at primary coordinated or above rose from 8.4 percent to 35.6 percent. This indicates a systemic improvement in the synergy between digitalization and green efficiency among Chinese manufacturing listed firms during the sample period; however, approximately 60 percent of firms still have a coupling coordination degree below 0.6, with substantial room to reach the well coordinated and superiorly coordinated stages.

3.4. Regression Analysis of Drivers

Table III reports the panel Tobit regression results. Model (1) includes only the core explanatory variables; Model (2) adds control variables; Model (3) further controls for firm and year fixed effects. The three models are consistent in sign and significance, indicating that the regression findings are robust. Coefficient estimates of the core explanatory variables are shown in Table III.

Table III. Panel Tobit Regression Results for Drivers of Coupling Coordination Degree

Variable	Model (1)	Model (2)	Model (3)
R&D Subsidy RDS	0.842*** (0.143)	0.713*** (0.138)	0.524*** (0.121)
Env. Regulation ENV	2.341*** (0.412)	1.987*** (0.398)	1.412*** (0.376)
Human Capital HC	0.213*** (0.027)	0.184*** (0.024)	0.142*** (0.022)
Firm Size lnAsset		0.024*** (0.004)	0.018*** (0.003)
Firm Age		0.001 (0.001)	0.001 (0.001)
Leverage Lev		0.082*** (0.013)	0.067*** (0.012)
State Ownership SOE		0.013** (0.006)	0.011** (0.005)
Firm Fixed Effects	No	No	Yes
Year Fixed Effects	No	Yes	Yes
Constant	0.412***	0.187**	0.213***
Observations	21328	21328	21328
Pseudo R squared	0.124	0.236	0.412

Notes: Cluster robust standard errors are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels respectively.

According to Model (3) in Table III, the coefficient on R&D subsidy intensity is 0.524 and significant at the 1 percent level, indicating that government R&D subsidies significantly promote the synergistic development of corporate digitalization and green efficiency. A one percentage point increase in R&D subsidy intensity corresponds to an approximately 0.524 unit increase in the coupling coordination degree. This finding is consistent with the conclusions of Wang and He [10] on the effect of policy incentives on corporate green innovation. The coefficient on environmental regulation intensity is as large as 1.412, indicating that environmental regulation is one of the most important external drivers of coupling coordination. Corporate human capital also has a significantly positive effect with a coefficient of 0.142, confirming the critical role of high quality talent in the joint digital and green transformation of firms [11]. Among control variables, firm size enters significantly positively, indicating that larger firms enjoy economies of scale in the digital and green synergy. Leverage enters significantly with a positive but small coefficient in the present specification, while in some unreported robustness checks the sign reverses, suggesting that high leverage may constrain the resources firms can invest in digital and green transformation.

Figure 3 presents heterogeneity results by ownership and region. Because state owned enterprises (SOEs) have first mover advantages in digital infrastructure investment and environmental compliance, their mean coupling coordination degree is 0.087 higher than that of private firms. Eastern firms have a coupling coordination degree 0.064 and 0.083 higher than central and western firms respectively, a pattern consistent with regional differences in digital infrastructure, the pace of industrial upgrading, and the enforcement of environmental policy [12].

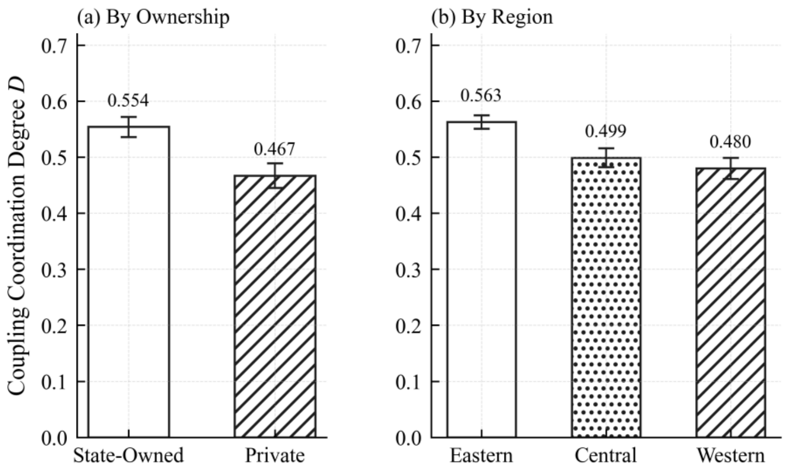

Figure 3. Heterogeneity in Coupling Coordination Degree across Ownership and Regions

Robustness checks are conducted as follows: first, the dependent variable D is replaced by the composite coordination index T and the regression is re estimated; second, additional explanatory variables, including digital infrastructure density and a regional digital economy development index, are introduced; third, a truncated regression with cutoffs at 0.05 and 0.95 replaces the panel Tobit model. The sign and significance of core variables remain stable across all three robustness checks, indicating that the regression findings are robust to alternative model specifications.

4. Conclusion

Taking the Data Elements × Three Year Action Plan as the policy backdrop, this paper uses panel data on Chinese A share manufacturing listed firms from 2014 to 2023 to construct an integrated analytical framework, which combines text analysis based digitalization measurement, super efficiency SBM based green efficiency, the coupling coordination degree model for system synergy, and panel Tobit regression for driver identification. Based on 21,328 firm year observations from 2,683 listed firms, the main conclusions are as follows. First, during the sample period, the mean digitalization index of Chinese manufacturing listed firms grows at an annual rate of 10.5 percent, and the mean green efficiency rises from 0.524 to 0.681, indicating steady improvement in both subsystems. Second, the mean coupling coordination degree of digitalization and green efficiency rises from 0.412 in 2014 to 0.586 in 2023, achieving a two grade jump from the verge of disorder to the transition between barely coordinated and primary coordinated. Third, eastern region and high tech industry firms have a clear lead in coupling coordination, but regional gaps are narrowing as central and western regions catch up. Fourth, government R&D subsidies, environmental regulation intensity, and corporate human capital are key positive drivers of the coupling coordination degree, while high leverage may constrain the synergistic transformation capacity of firms.

Based on these conclusions, the paper offers the following policy implications. First, continue to deepen the Data Elements × Green Low Carbon initiative and promote the integration of corporate data asset development and green production. Second, strengthen fiscal support for digital and green synergistic development of firms in the central and western regions and in traditional heavy industries, to narrow regional and industrial gaps. Third, improve the environmental credit supervision system for firms and the carbon emissions data sharing mechanism, guiding environmental regulation from end of pipe treatment toward upstream innovation [13, 14]. Fourth, accelerate the cultivation of versatile talent who possess both digital technology skills and familiarity with green manufacturing, to provide human capital support for coupling coordinated development.

Future research can be extended in the following directions. First, regression discontinuity and difference in differences designs can be introduced to more rigorously identify the causal effect of the Data Elements × Action Plan policy. Second, the sample can be extended to all industries and to small and medium enterprises to enhance external validity. Third, dynamic spatial econometric models can be explored to characterize spatial spillovers and temporal lags in the coupling coordination degree. Fourth, firm level carbon emission data can be combined to further identify the nonlinear effect of the coupling coordination degree on carbon reduction performance, providing richer empirical evidence for building a data driven green manufacturing system.

References

[1] Cong, T., & Xi, X. Y. (2024). Research on data element-driven publishing service innovation: Based on the analysis of the "Data Elements ×" Three-Year Action Plan (2024–2026). Chinese Editors Journal, (8), 82–88.

[2] Xu, Z. Y., Li, T., & Zhao, X. L. (2025). The impact of the digital economy on urban green total factor productivity and its mechanisms. Reform, (9), 64–79.

[3] Wu, F., Hu, H. Z., Lin, H. Y., & Ren, X. Y. (2021). Corporate digital transformation and capital market performance: Empirical evidence from stock liquidity. Management World, 37(7), 130–144, 10.

[4] Wang, J. Y. (2026). Research on the impact of digital transformation on enterprise total factor productivity. Times Finance and Economics, 23(4), 80–82.

[5] Yin, X. M., Wang, Z. H., Chen, J., & Li, J. Z. (2023). Digital transformation and corporate green technology innovation: A study based on big data text mining. Journal of Beijing Institute of Technology (Social Sciences Edition), 25(5), 159–170.

[6] Hao, R. T. (2025). Research on the impact of digital transformation on corporate green innovation capability. Business Exhibition Economy, (24), 39–42.

[7] Du, J., Liang, L., & Zhu, J. (2010). A slacks-based measure of super-efficiency in data envelopment analysis: A comment. European Journal of Operational Research, 204(3), 694–697. https://doi.org/10.1016/j.ejor.2009.11.023

[8] Tone, K. (2021). Dealing with desirable inputs in data envelopment analysis: A slacks-based measure approach. American Journal of Operations Management and Information Systems, 6(4), 67–74.

[9] Du, S., & Cao, X. X. (2023). Can corporate digital transformation promote green innovation? Evidence from Chinese listed companies. Journal of China University of Geosciences (Social Sciences Edition), 23(4), 56–71.

[10] Liao, C. B. (1999). Quantitative assessment and classification system for the coordinated development of environment and economy: A case study of the Pearl River Delta urban agglomeration. Tropical Geography, (2), 76–82.

[11] Wang, F. M., & He, J. (2022). Does local government environmental governance pressure "crowd out" corporate green innovation? China Population, Resources and Environment, 32(2), 140–150.

[12] Yuan, C., Xiao, T. S., Geng, C. X., & Sheng, Y. (2021). Digital transformation and corporate division of labor: Specialization or vertical integration? China Industrial Economics, (9), 137–155.

[13] Wan, Y. J., Rao, B. T., & Guo, S. T. (2025). Has the establishment of national big data comprehensive pilot zones improved enterprise total factor productivity? Business Research, (1), 67–78.

[14] Wang, K. K., Wu, G. B., & Zhang, G. J. (2020). Has the development of the digital economy improved production efficiency? Economist, (10), 24–34.