Data Element Marketization and the Shift in Green Innovation: From Symbolic to Substantive

Jiadong Wang, Yinghui Ma *, Jiayi Xu, Zhilin Fan

School of Economics and Management, Jiangsu University, Zhenjiang, China, 212100

* Corresponding Author Email: myh@just.edu.cn

Abstract. Promoting substantive green innovation is critical for sustainable economic growth. However, the “double externality” of environmental protection often incentivizes firms to prioritize symbolic innovation for image management over high-quality substantive innovation. This paper investigates whether Data Element Marketization (DEM) can correct this distortion and optimize the structure of corporate green innovation. Treating the establishment of data trading platforms as a quasi-natural experiment, this study employs a staggered difference-in-differences (DID) design using a sample of Chinese A-share listed companies from 2009 to 2023. By differentiating between substantive and symbolic innovation, the research rigorously tests the impact of DEM. The empirical results demonstrate that DEM significantly promotes substantive green innovation while having a limited and statistically insignificant impact on symbolic innovation. Mechanism analysis reveals that DEM drives this shift by reducing information asymmetry, alleviating financing constraints and enhancing management efficiency. Furthermore, heterogeneity analysis indicates that the positive effect is more pronounced in high-tech industries, firms with lower environmental awareness, and regions with under-developed digital finance. This study contributes to the literature by uncovering the unique role of DEM in enterprise green innovation. The findings suggest that policymakers should integrate DEM into environmental governance frameworks to encourage firms to move from symbolic to substantive environmental advancements.

Keywords: Data Element Marketization, Green Innovation Structure, Substantive Innovation, Symbolic Innovation.

1. Introduction

Sustainable development is a global imperative, and green innovation is essential for environmental protection and economic growth. Green innovation addresses environmental challenges, fostering sustainable economic growth through enhanced energy efficiency and optimized resource use [1]. However, information asymmetry and market failure significantly hinder green innovation [2]. This "double externality" distorts incentives, leading firms to favor symbolic innovations over substantive innovations[3] . Unlike substantive innovations, which offer direct and substantial environmental benefits, symbolic innovations are often focused on image management and may not lead to significant environmental improvements. Correcting this imbalance is essential to better match corporate behavior with sustainable development goals [4]. DEM may promote substantive green innovation by enhancing information flow and reducing transaction costs.

Prior research on green innovation has largely focused on traditional factors like environmental regulations and market demand [5,1]. While some studies examine the digital economy's role [6], they often overlook the specific contribution of DEM in fostering green innovation. Unlike traditional markets, data element markets, often facilitated by platforms or broader data access, can reduce information asymmetry [7]. Lowering the cost of searching for innovation-relevant data allows firms to optimize resource use and advance technology. The real-time nature of data traded on these platforms further enhances green innovation responsiveness [8]. Therefore, DEM uniquely transforms data into a dynamic input for innovation. This distinct mechanism necessitates further investigation into its specific impact on substantive green innovation.

This study contributes to the literature by: 1) expanding the understanding of DEM's novel role in driving corporate innovation, especially substantive green innovation; 2) advancing the green innovation literature by differentiating between substantive and symbolic green innovation and demonstrating that DEM specifically fosters the former; and 3) uncovering the underlying mechanisms through which DEM impacts substantive green innovation using firm-level data.

2. Hypothesis development

We propose that DEM optimizes corporate green innovation structure, particularly by increasing substantive innovation. DEM primarily achieves this outcome by reducing information asymmetry. Data available in the market accurately reflects a company's operational and environmental performance, increasing transparency and enabling stakeholders to access more reliable information [9]. This reduces asymmetry and encourages firms to focus on long-term, substantive innovation [10]. A transparent information environment is essential for stakeholders to effectively evaluate and monitor corporate environmental performance. Furthermore, DEM mitigates financing constraints, a crucial factor for green innovation. By enabling data elements to serve as assets, marketization broadens financing channels, particularly for high-risk, long-cycle substantive green innovation projects [11]. This broadening is achieved through enhanced information flows and improved management practices. As a result, firms can better secure external funding [9]. The alleviation of such constraints is particularly critical for substantive green innovation, given its frequent requirement for substantial capital investment. Finally, DEM enhances management efficiency and optimizes resource allocation. Access to richer market and technological information facilitated by marketization significantly improves the quality of decision-making [12]. It promotes more efficient internal resource allocation, thereby improving overall resource utilization. In a data-driven market environment, corporate environmental performance emerges as a key evaluation metric, consequently motivating firms to prioritize substantive green innovation for sustained long-term benefits[13]. Therefore, we propose the following hypothesis:

Hypothesis 1: DEM optimizes corporate green innovation structure by increasing substantive innovation.

Hypothesis 2: DEM reduces information asymmetry, alleviates financing constraints, and enhances management efficiency, leading to an increase in substantive green innovation.

3. Data and methods

This study uses a sample of Chinese A-share listed companies from 2009 to 2023. Corporate green innovation (Total_GI) is measured using the number of green patent applications. Substantive and symbolic green innovation(Substantive_GI and Symbolic_GI) are measured by the number of invention and utility model patents applications, respectively (Li, Yu, & Qian, 2024; Li & Zheng, 2016). To quantify the degree of DEM, this study uses the establishment of data trading platforms as a proxy. We obtained data on green innovation from the CNRDS database, and corporate financial data from the CSMAR database. After excluding the financial industry and firms marked ST, *ST, or PT, the final dataset comprises 31861 firm-year observations. All the variables are winsorized at 1%.

To examine the impact of DEM on corporate green innovation structure, we employ the following staggered difference-in-differences (DID) design:

$$Y_{ist} = \beta_0 + \beta_1 DEM_{ist} + \gamma X_{ist} + \delta_i + \lambda_t + \mu_{st} + \varepsilon_{ist} \tag{1}$$

where i denotes firms and t represents years. All variable definitions and descriptive statistics are provided in Table 1.

Table 1. Descriptive Statistics

	N	Mean	Std	Min	Max	Variable definitions
Total_GI	31861	0.303	0.711	0.000	8.318	Natural logarithm of one plus a firm’s total green patent applications.
Substantive _GI	31861	0.254	0.619	0.000	6.931	Natural logarithm of one plus a firm’s green utility patent applications.
Symbolic_GI	31861	0.436	0.875	0.000	10.986	Natural logarithm of one plus a firm's green invention patent applications.
Age	31861	1.352	0.095	0.741	1.528	Natural logarithm of one plus the number of establishment years.
Lev	31861	0.421	0.202	0.028	1.262	Ratio of total debt to total assets.
Roa	31861	0.037	0.070	-0.703	0.366	Ratio of net profit to total assets.
Soe	31861	0.373	0.484	0.000	1.000	A dummy variable equal to 1 for state-owned firms, 0 otherwise.
Indep	31861	0.375	0.054	0.250	0.600	Ratio of independent directors to total directors.
Top1	31861	0.344	0.147	0.079	0.753	Ratio of shares held by the largest shareholder to outstanding shares.
Dual	31861	0.274	0.446	0.000	1.000	A dummy for CEO duality (1 if CEO is also Chairperson, 0 otherwise).
Growth	31861	0.176	0.446	0.759	5.491	Annual revenue growth rate of a firm.
Green_CEO	31861	0.092	0.289	0.000	1.000	A dummy for CEO's green background (1 if yes, 0 otherwise).
Note: This table presents descriptive statistics for the main variables used in our analysis.						

4. Empirical Results

4.1. Baseline Results

Table 2 presents the baseline regression results for the impact of DEM on the structure of corporate green innovation. The results reveal that the coefficient of DEM is significantly positive at the 1% level in the regressions for Total_GI in columns (1) and (2) and in the regressions for Substantive_GI in columns (3) and (4). This finding indicates that DEM significantly promotes overall Total_GI, particularly in Substantive_GI. However, the coefficient of DEM is not significant in the regressions for Symbolic_GI in columns (5) and (6). Taken together, these results suggest that DEM optimizes the structure of corporate green innovation, shifting the focus toward higher-quality substantive innovation rather than merely symbolic innovation activities.

Table 2. Baseline regression results

	(1)	(2)	(3)	(4)	(5)	(6)
	Total_GI	Total_GI	Substantive _GI	Substantive _GI	Symbolic_GI	Symbolic_GI
DEM	0.057^{***}	0.059^{***}	0.064^{***}	0.064^{***}	0.024	0.026
	(0.020)	(0.020)	(0.016)	(0.016)	(0.024)	(0.024)
Control	N	Y	N	Y	N	Y
Firm FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Province×Year FE	Y	Y	Y	Y	Y	Y
N	31861	31861	31861	31861	31861	31861
R^2	0.643	0.643	0.625	0.625	0.577	0.578
Notes: (1) Standard errors clustered at the firm level are reported in parentheses. (2) ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.						

4.2. Dynamic Effects

We use an event study framework to investigate the dynamic impact of DEM on Substantive_GI (see equation below). I_j is a dummy for the relative period, ranging from -4 to 8, excluding the year prior to the event (-1):

$$Y_{ist} = \beta_0 + \sum_{j=-4, j\neq -1} \beta_j \, DEM_i \times I_j + \gamma X_{ist} + \delta_i + \lambda_t + \mu_{st} + \varepsilon_{ist} \quad (2)$$

The results are illustrated in Figure 1. Prior to the policy event (year 0), the coefficients of the interaction terms fluctuated around zero, suggesting that the treatment group and control group satisfied the parallel trends assumption ex-ante. After the policy event, the coefficients exhibit a clear upward trend and become significantly positive, indicating that DEM has a positive and increasing impact on Substantive_GI.

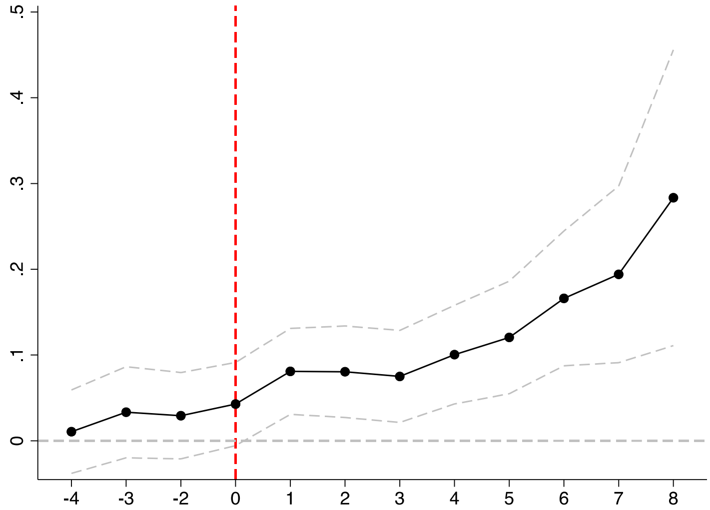

Figure 1. The dynamic effects of policy.

4.3. Placebo Test

We conducted a placebo test, randomly sampling the interaction term 500 times. Figure 2 shows the simulated coefficients clustered around zero, significantly smaller than our baseline coefficient (0.064), confirming our results are not due to chance.

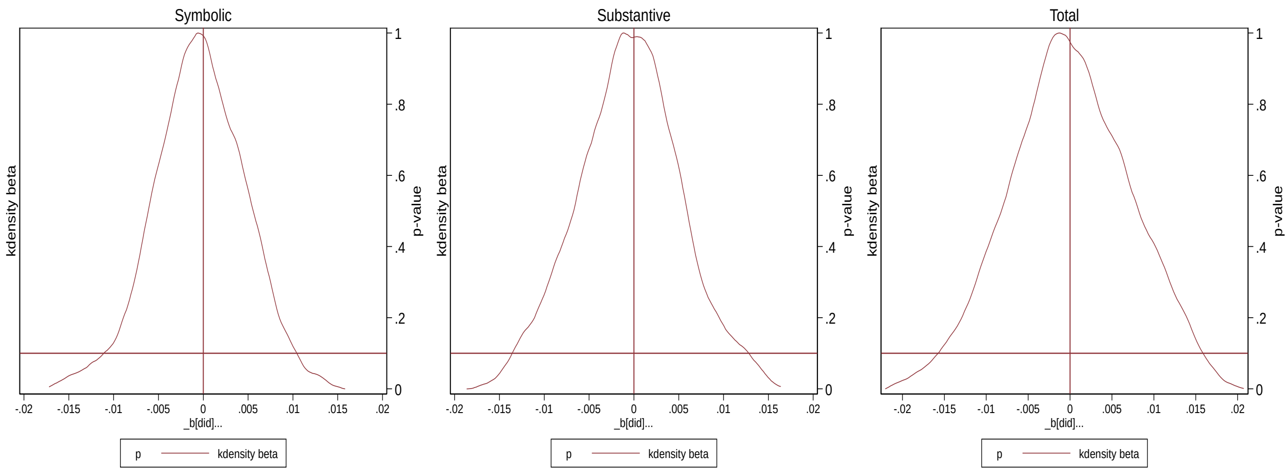

Figure 2. Placebo test

4.4. Additional Robustness Tests

We conducted several supplementary analyses to further ensure robustness, including PSM-DID to address selection bias, controlling for other policy interventions, using an alternative dependent variable, excluding 2020 due to the COVID-19 pandemic, testing for lagged effects, and comparing FE and RE models. As shown in Table 3 (columns 1-7), the estimated coefficients on DEM remain significant ($p < 0.05$) and consistent with the baseline results, confirming robustness.

Table 3. Robustness Test Results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	PSM-DID	Other Policy	Alt. DV	No 2020	Lag Effect	FE	RE
DEM	0.063***	0.056***	0.048***	0.071***	0.039**	0.054***	0.071***
	(0.017)	(0.017)	(0.012)	(0.017)	(0.017)	(0.018)	(0.016)
Carbon		0.022					
		(0.042)					
Band		0.060***					
		(0.020)					
Control	Y	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	N	N
Year FE	Y	Y	Y	Y	Y	Y	Y
Prov×Year FE	Y	Y	Y	Y	Y	N	N
N	27756	27544	28075	28672	27756	31861	31861
R^2	0.702	0.702	0.649	0.674	0.679	0.022	-
Notes: (1) Standard errors clustered at the firm level are reported in parentheses. (2) ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.							

5. Further discussion

5.1. Heterogeneity analysis

Table 4 presents the heterogeneity analysis, examining how the impact of DEM on corporate green innovation varies across firm and regional characteristics, including environmental awareness (EA), high-tech industry (HTI) status, and digital inclusive finance (DIF) levels. DEM has a more pronounced positive effect on Substantive_GI in firms with low EA (column 1), high-tech industries (column 4), and regions with lower DIF (column 5). The Chow test confirms that the coefficients are significantly different across these groups, supporting the presence of heterogeneous effects.

Table 4. Heterogeneity analysis.

	EA		HTI		DIF	
	(1)	(2)	(3)	(4)	(5)	(6)
	Low	High	Low	High	Low	High
DEM	0.078***	0.032	0.033	0.091***	0.095***	0.039
	(0.026)	(0.046)	(0.020)	(0.024)	(0.021)	(0.030)
Control	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Prov×Year FE	Y	Y	Y	Y	Y	Y
N	16498	14622	12669	19150	14482	17070
R^2	0.599	0.676	0.587	0.630	0.693	0.647
Notes: (1) Standard errors clustered at the firm level are reported in parentheses. (2) ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.						

5.2. Mechanism analysis

Table 5 examines the pathways through which DEM affects corporate green innovation, using three intermediary variables: information asymmetry (ASY), financing constraints (FC), and management efficiency (ME). DEM significantly reduces information asymmetry (ASY, column 1), financing constraints (FC, column 2), and management costs (column 3, proxying for management efficiency). Thus, DEM optimizes green innovation through the mechanisms of reducing information asymmetry, alleviating financing constraints, and improving management efficiency.

Table 5. Mechanism Analysis Results

	(1)	(2)	(3)
	ASY	FC	ME
DEM	-0.020**	-0.011***	-0.009***
	(0.009)	(0.004)	(0.002)
Control	Y	Y	Y
Firm FE	Y	Y	Y
Year FE	Y	Y	Y
Prov×Year FE	Y	Y	Y
N	31861	31861	31861
R^2	0.686	0.822	0.658
Notes: (1) Standard errors clustered at the firm level are reported in parentheses. (2) ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.			

6. Conclusions

Using a quasi-natural experiment of data trading platforms in China, this study reveals that DEM significantly optimizes corporate green innovation structure by promoting substantive over symbolic forms. This effect operates through reduced information asymmetry, alleviated financing constraints, and enhanced management efficiency. These findings suggest that policymakers should integrate data market development into environmental governance frameworks, while business leaders can leverage data markets to enhance substantive green innovation capabilities. This study advances our understanding of how digital transformation can drive genuine substantive environmental initiatives.

Acknowledgments

This research was supported by the humanities and social science youth project funded by the Ministry of Education of People's Republic of China (21YJC790086) and the National College Students' Innovation and Entrepreneurship Training Program of Jiangsu University of Science and Technology (202510289039).

References

[1] Chen Y S, Lai S B, Wen C T. The influence of green innovation performance on corporate advantage in Taiwan [J]. Journal of Business Ethics, 2006, 67 (4): 331-339.

[2] Rennings K. Redefining innovation — eco-innovation research and the contribution from ecological economics [J]. Ecological Economics, 2000, 32 (2): 319-332.

[3] Jiang L, Bai Y. Strategic or substantive innovation? -The impact of institutional investors' site visits on green innovation evidence from China [J]. Technology in Society, 2022, 68: 101904.

[4] Sun Z, Sun X, Wang W, et al. Digital transformation and greenwashing in environmental, social, and governance disclosure: Does investor attention matter? [J]. Business Ethics, the Environment &Amp; Responsibility, 2023, 34 (1).

[5] Porter M E, Van Der Linde C. Toward a new conception of the environment-competitiveness relationship[M]//Economic Costs and Consequences of Environmental Regulation. 2018: 413-434.

[6] Miao C lin, Duan M meng, Zuo Y. Spatial heterogeneity and evolution trend of regional green innovation efficiency–an empirical study based on panel data of industrial enterprises in China’s provinces [J]. Energy Policy, 2021, 156: 112370.

[7] Hagiu A, Wright J. Multi-sided platforms [J]. International Journal of Industrial Organization, 2015, 43: 162-174.

[8] Gao S, Li W, Meng J, et al. A Study on the Impact Mechanism of Digitalization on Corporate Green Innovation [J]. Sustainability, 2023, 15 (8).

[9] Fernandez R C, Subramaniam P, Franklin M J. Data Market Platforms: Trading Data Assets to Solve Data Problems [M]. arXiv, 2020.

[10] Cuypers I, Koh P S, Wang H. Sincerity in Corporate Philanthropy, Stakeholder Perceptions and Firm Value [J]. Organization Science, 2016, 27 (1).

[11] Altaira H, Margaretha Leon F. Effect of Green Financing and Financing Constraints on Green Technology Innovation in the Consumer Goods Industry in Indonesia [J]. International Journal of Management Studies and Social Science Research, 2023, 05 (01).

[12] Franke F, Hiebl M R W. Big data and decision quality: the role of management accountants’ data analytics skills [J]. International Journal of Accounting & Information Management, 2022, 31 (1): 93-127.

[13] Liu M, Liu L, Feng A. The Impact of Green Innovation on Corporate Performance: An Analysis Based on Substantive and Strategic Green Innovations [J]. Sustainability, 2024, 16 (6): 2588.