Article

Not peer-reviewed version

The Impact of Open Public Data on Corporate Low-Carbon Technological Innovation: Evidence from China

Jing Wang , Jie Wang , Zhijian Cai *

Posted Date: 4 November 2025

doi: 10.20944/preprints202511.0268.v1

Keywords: open public data; low-carbon technological innovation; government transparency; barriers to factor mobility; spillover effects

Article

The Impact of Open Public Data on Corporate Low-Carbon Technological Innovation: Evidence from China

Jing Wang [1], Jie Wang [2] and Zhijian Cai [1,*]

[1] College of Economics and Management, Nanjing Forestry University, Nanjing 210037, China

[2] Business School, Chizhou University, Chizhou 247000, China

* Correspondence: janecai69@njfu.edu.cn

Abstract

Open public data is a vital institutional arrangement for overcoming data constraints in corporate low-carbon technological innovation; however, its mechanisms and boundaries remain empirically untested. Using a panel dataset of China's Shanghai- and Shenzhen-listed A-share firms over the 2007-2023 period, this study employs a difference-in-differences (DID) approach to examine the impact of open public data on corporate low-carbon technological innovation. The results show that open public data has a significant positive effect on corporate low-carbon technological innovation, and the results remain robust across multiple validation tests. Mechanism tests point out that open public data primarily drives corporate low-carbon technological innovation by enhancing government transparency and reducing barriers to factor mobility. The heterogeneity analysis indicates that the positive impact of open public data is more pronounced among firms characterized by higher R&D investment, lower financial constraints, and greater digitalization. Further analysis indicates that open public data also exhibits significant geographic and industry spillover effects, with the geographic spillover following an inverted U-shaped pattern of decay and the industry spillover driven by peer imitation. This paper provides policy references for the development of open public data and the enhancement of corporate low-carbon technological innovation.

Keywords: open public data; low-carbon technological innovation; government transparency; barriers to factor mobility; spillover effects

---

1. Introduction

The 2025 Sustainable Development Goals Report indicates that the global climate crisis is rapidly worsening, with temperatures approximately 1.55°C above pre-industrial levels and global greenhouse gas emissions reaching record levels in 2023. This grim reality serves as a stark warning to the world to take decisive action. As the world's largest energy consumer and carbon dioxide emitter, China's emissions reduction pathway holds critical implications for global climate governance [1]. To this end, China has proactively set targets to attain a "carbon peaking" by 2030 and "carbon neutrality" by 2060. Low-carbon technological innovation is pivotal to achieving these goals, encompassing innovations in renewable energy, energy efficiency improvements, and carbon capture and storage technologies [2]. However, corporate low-carbon technological innovation faces two major barriers: limited access to relevant data resources and insufficient capacity to integrate advanced information into innovation [3]. Therefore, researching how to effectively incentivize and empower enterprises to innovate in low-carbon technologies has become a critical practical issue.

This paper examines the effects of open public data on promoting corporate low-carbon technological innovation. In today's digital economy, data function as more than just a new production input—they have become a vital force driving both economic growth and social advancement [4]. In 2024, China's total data output reached 41.06 zettabytes, representing a 25% year-

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 November 2025 doi:10.20944/preprints202511.0268.v1

on-year increase[1]. Public data constitutes the largest share of data resources and plays a foundational, leading, and exemplary role in the development and utilization of data elements. Open public data can dismantle "information silos," mitigate corporate data insufficiency, reduce information search costs, optimize factor allocation, alleviate information asymmetry between government and enterprises [5], and ultimately enhance low-carbon technological innovation.

Open public data refers to government efforts to integrate public data resources and make them universally accessible, thereby safeguarding citizens' rights to access and use data and fostering the social and commercial value of data [6]. Prior studies have affirmed the institutional role of open public data in fostering corporate innovation [7,8]. Moreover, existing studies on the determinants of low-carbon technological innovation have mainly explored two pathways: firm-level resources and policy interventions. The resource-based perspective highlights the roles of internal R&D investment [9], managerial capability [10], and organizational adaptability [11], whereas policy-oriented studies emphasize the incentive effects of credit policies [9], carbon taxation [12], and pilot programs [13,14]. However, limited attention has been paid to how open public data alleviates bottlenecks in corporate low-carbon technological innovation or whether its influence generates spillover effects.

This study employs data from Chinese A-share listed firms on the Shanghai and Shenzhen stock exchanges from 2007 to 2023, and constructs a multi-period difference-in-differences model to empirically test the impact of public data openness on corporate low-carbon technological innovation. The empirical analysis reveals that open public data significantly promotes corporate low-carbon technological innovation, and this conclusion passes a series of robustness tests. Furthermore, the study explores the mechanisms through which open public data affects low-carbon technological innovation and assesses how firm-level heterogeneity shapes its effectiveness. Further analysis reveals that the positive impact of open public data on corporate low-carbon technological innovation exhibits spillover effects at both the geographic and industry levels.

The primary contributions of this study are threefold. First, this study enriches empirical research on open public data at the enterprise level. By extending the microeconomic analysis of open public data to the context of sustainable and low-carbon transition, this study provides the first systematic evaluation of its role in driving corporate low-carbon innovation. Our findings offer direct firm-level evidence for the integrated "data factor × green development" strategy. Second, it elucidates the mechanisms by which open public data affects corporate low-carbon technological innovation. Moving beyond the conventional notion that "openness inherently creates value", the empirical evidence reveals that public data openness drives low-carbon technological innovation of enterprises through two key paths, namely, enhancing government transparency and lowering the barriers of factor circulation, which deepens the theoretical understanding of how data elements can be transformed into the driving force of green innovation. Finally, this research provides scientific evidence for advancing data openness from universal provision to targeted collaboration. Findings reveal that policy effectiveness is influenced by firm-level factors (R&D, financing, digitalization) and exhibits geographic decay boundaries alongside industry clustering effects. This offers crucial empirical support for policymakers to shift from a one-size-fits-all approach toward differentiated empowerment and the construction of a cross-domain collaborative data ecosystem.

The rest of the paper is structured as follows: Section 2 outlines the institutional background and theoretical hypotheses. Section 3 describes research methodology. Section 4 presents specific empirical results and analysis. Section 5 provides further discussion. Section 6 summarizes conclusions, policy implications, and limitations.

---

[1] China Academy of Information and Communications Technology (CAICT) and China Cyberspace Research Institute, National Data Resources Survey Report (2024), April 2025.

2. Institutional Background and Theoretical Hypotheses

2.1. Policy Background

The Chinese government places significant emphasis on institutionalizing public data resources. In 2015, the release of the Outline of Action for Big Data Development elevated government data openness to a national strategy for the first time, explicitly stating that "government data openness should drive the circulation of social data." This high-level design has been further developed in subsequent policies. In 2022, the 14th Five-Year Plan for the Development of the Digital Economy further proposed establishing a national public data resource system, emphasizing data openness to resolve resource mismatches. In the same year, the Opinions on Building a Data Infrastructure System to Better Leverage the Role of Data as a Production Factor introduced an innovative mechanism for authorizing public data rights, fostering a shift in data property rights from "closed and exclusive" to "open and shared." With the launch of policies such as the Three-Year Action Plan for "Data Element ×" (2024–2026), expanding public data supply and promoting well-regulated data openness have become key priorities in China's data sector in recent years. Over the past decade, the Chinese government has actively promoted the development of local open data platforms. In 2012, only two provincial-level administrative regions spearheaded pilot programs; by 2024, that number had grown to 28 provinces with established platforms[2].

2.2. Theoretical Hypotheses

Open public data can alleviate data constraints on enterprises, thereby fostering low-carbon technological innovation. This is because: First, the public data released by the government is a "public good". Enterprises can access, utilize, and analyze the knowledge, market and policy information contained within this data at little or no cost [15]. This allows enterprises to more effectively pinpoint the strategic direction of their low-carbon innovation efforts. Moreover, public data is authoritative and reliable. Enterprises that access these data resources can reduce unnecessary costs associated with data search, verification, and trial-and-error [16,17], thereby effectively reducing the foundational costs of low-carbon technological innovation. It can be seen that open public data improves the generalizability of data resources and lowers the threshold of access to them, thereby breaking down information barriers faced by enterprises to a certain extent [6]. Reducing information barriers enables firms to better discern policy directions and reallocate resources toward production and R&D activities [18], thereby facilitating low-carbon technological innovation. Based on the above analysis, this study proposes the following hypothesis.

H1. Open public data promotes corporate low-carbon technological innovation.

On the one hand, open public data can enhance government transparency and alleviate information asymmetry between governments and firms, thereby promoting corporate innovation in low-carbon technologies. Due to high collection costs, centralized ownership, and strong administrative attributes [19], public data inherently exhibits monopolistic characteristics in its sources. This exacerbates information asymmetry between government and enterprise and impedes governmental transparency. A lack of transparency may contribute to strategic myopia among managers, thereby weakening their willingness to pursue long-term low-carbon innovation [20]. Open public data significantly enhances enterprises' access to public-sector information, enabling them to better understand government decision-making processes and policy orientations, and to build and strengthen accountability mechanisms for public supervision and participation, thereby effectively enhancing government transparency [21]. As a result, improved government transparency optimizes the institutional environment, mitigates the "lemons market" problem [22] in low-carbon

---

[2] Lab for Digital & Mobile Governance of Fudan University. (2024). China local government data opening report (2024).

 doi:10.20944/preprints202511.0268.v1

technological innovation, and reduces the decision-making risks enterprises face due to policy uncertainty, thereby fostering a more stable and predictable environment for corporate low-carbon technological innovation.

On the other hand, open public data can dismantle barriers to factor mobility, optimizing distribution in the data factor market, and thereby promoting corporate low-carbon technological innovation. Large companies that already possess vast amounts of data and advanced technologies enjoy significant cost advantages in acquiring new data [23]. They can leverage existing data assets and analytical capabilities to extract value more efficiently from new data sources, thereby creating a sustained competitive advantage [24,25]. This dynamic reinforces market concentration and exacerbates a fundamental competitive asymmetry between data-rich and data-poor entities. This dynamic results in a new form of asymmetry among market participants—the widening gap between “data-rich” entities (firms possessing extensive data resources and advanced analytical capabilities) and “data-poor” entities (small and medium-sized enterprises or individuals), placing the latter at a significant disadvantage in decision-making. Open public data helps correct this imbalance by lowering entry barriers and allowing broader access to valuable information. This enables knowledge and data to flow more freely—particularly benefiting small and medium-sized innovators with limited resources—and thus stimulates low-carbon technological innovation across the broader market.

Building on this theoretical reasoning, this paper proposes the following hypothesis.

H2. Open public data promotes corporate low-carbon technological innovation by enhancing government transparency and dismantling barriers to factor mobility.

3. Research Design

3.1. Sample Selection and Data Sources

This study examines A-share-listed firms on China’s Shanghai and Shenzhen stock exchanges from 2007 to 2023 as the research sample. To enhance the representativeness of the sample, the study employs the following approach: exclusion of (1) ST, *ST, and PT firms; (2) financial and insurance firms; and (3) observations with significant missing data. Finally, we obtained 37705 firm-year observations. To mitigate the impact of outliers on regression results, all continuous variables were trimmed at the 1% and 99% tails. Regarding data sources, the firm-level data are sourced from the CSMAR database, and patent data are retrieved from the China National Intellectual Property Administration and the Global Patent Database. The data on provincial open public data platforms are derived from the China Local Government Data Openness Report, an annual publication issued by the Digital and Mobile Governance Lab of Fudan University.

3.2. Variable Construction

3.2.1. Dependent Variable

Corporate low-carbon technological innovation (Lct). While patent data is widely adopted in existing research as a quantitative metric, this approach suffers from significant limitations. Green patents are broader in scope, covering various aspects of environmental protection and sustainable development. In contrast, low-carbon patents concentrate on innovations explicitly designed to support low-carbon technological advancement. To enhance measurement precision, this study draws on Zhu et al. [26] and Chen et al. [9], selecting patents from two primary categories closely related to low-carbon technologies (i.e., alternative energy production and energy-saving technology patents) based on the IPC Green List classification standards. These patents were aggregated, log-transformed, and incorporated into the model.

3.2.2. Core Independent Variable

Open public data ($Open$). Open public data is a dummy variable indicating whether the province where a firm is registered has launched a public data platform. It takes the value 1 in the year the platform is launched and in all subsequent years, and 0 otherwise.

3.2.3. Core Independent Variable

Referring to prior research, this paper incorporates a set of firm-level characteristics as control variables: gearing ratio (Lev), return on assets (Roa), firm size ($Size$), firm age (Age), social wealth creativity ($Tobinq$), board size ($Board$), the ratio of management shareholding ($Mhold$), and independent directors (Ind). Table 1 reports summary statistics for all variables.

Table 1. Descriptive statistics.

Variables	Definition	Obs.	Mean	SD
Lct	Logarithm of the number of low-carbon patent applications	37705	0.194	0.553
Open	A value of 1 is assigned when the firm's province has implemented an open public data platform; otherwise, the value is 0.	37705	0.575	0.494
Lev	The proportion of total debts to total assets	37705	0.398	0.197
Roa	Net profit after tax to total assets	37705	0.044	0.054
Size	The natural logarithm of total assets	37705	22.155	1.271
Age	The natural logarithm of the current year minus the year of listing plus one.	37705	1.913	0.940
Tobinq	The ratio of firm market value to replacement capital	37705	1.963	1.118
Board	Number of directors on the board	37705	2.238	0.177
Mhold	The ratio of management shareholding	37705	0.153	0.205
Idr	The ratio of independent directors	37705	0.376	0.053

3.3. Estimation Model

Given the phased implementation of open public data, this study designates provinces that have launched public data platforms as the treatment group, with the remainder serving as the control group. We employ a staggered DID method to evaluate the policy's impact on corporate low-carbon technological innovation. The baseline regression specification is constructed as follows:

$$Lct_{it} = \beta_0 + \beta_1 Open_{it} + \beta_2 Control_{it} + \lambda_i + \mu_t + \varepsilon_{it} \quad (1)$$

where i and t denote firm i and year t, respectively. Lct_{it} represents the level of low-carbon technological innovation of firm i in year t. $Open_{it}$ indicates whether the province of a firm's registration has launched a public data platform. $Control_{it}$ denotes a set of control variables. λ_i and μ_t are individual (firm) and year fixed effects, respectively. ε_{it} is the random error term, with robust standard errors clustered at the firm level.

4. Results and Analysis

4.1. Benchmark Regression Results

Table 2 reports the results of the benchmark regression examining the impact of open public data on corporate low-carbon technological innovation. Column (1) displays baseline results with no control variables, whereas Column (2) presents the results incorporating the full set of covariates. It is discernible that the estimated coefficients for $Open$ in Columns (1) and (2) both demonstrate positive at the 1% significance level, confirming that open public data significantly promotes corporate low-carbon technological innovation. From an economic perspective, an increase of one standard deviation (0.494) in public data openness leads to an average 2.322% increase in corporate low-carbon technological innovation. Based on this, hypothesis 1 is validated.

 doi:10.20944/preprints202511.0268.v1

Table 2. Benchmark regression results.

Variables	Lct	
	(1)	(2)
Open	0.028***	0.026***
	(0.009)	(0.009)
Lev		0.015
		(0.032)
Roa		0.047
		(0.059)
Size		0.060***
		(0.010)
Age		0.011
		(0.010)
Tobinq		0.006**
		(0.003)
Board		-0.004
		(0.035)
Mhold		0.137***
		(0.034)
Idr		-0.020
		(0.087)
Firm FE	YES	YES
Year FE	YES	YES
Observations	37705	37705
R-squared	0.636	0.638

Notes: * p < 0.1, ** p < 0.05, *** p < 0.01, Standard errors in parentheses are clustered at the firm level.

4.2. Robustness Checks

4.2.1. Parallel Trends Test

Since the DID framework depends on the parallel trends assumption between treated and untreated firms, we follow Beck et al. [27] and disaggregate the event timeline around the policy implementation to examine the time-varying impact of public data openness on corporate low-carbon innovation. Additionally, to mitigate sample imbalance, observations from periods earlier than t = -5 are reset to -5, and those later than t = 4 are reset to 4, with t = -1 designated as the baseline (omitted) period. The dynamic DID model is specified as follows:

$$Lct_{it} = \beta_0 + \sum_{k=-5,k\neq-1}^{k=4} \beta_k \, Open_{it}^{k} + \beta_2 Control_{it} + \lambda_i + \mu_t + \varepsilon_{it} \qquad (2)$$

In the equation, $Open_{it}^{k}$ is a dummy variable indicating whether firm i is in the kth period of the pre-treatment phase ($k < 0$) or the post-treatment phase ($k \geq 0$) during time period t. The meanings of other variables remain consistent with baseline regression model (1).

The results of the parallel trends test are presented in Figure 1. At the condition of 95 % confidence interval, the coefficients $\beta_k(k < 0)$ fluctuate around 0 and are not significant before the establishment of the public data platform. This indicates that, prior to the platform launch, there was no significant difference in low-carbon technological innovation between treatment and control groups, satisfying the parallel trends assumption. Following the release of public data, the coefficients β_k ($k \geq 0$) became significantly positive and showed an upward trend, indicating that the promotional effect of open public data on corporate low-carbon technological innovation is both sustained and increasing. This suggests that the policy's effectiveness strengthens as the breadth and depth of data openness increase.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 November 2025 doi:10.20944/preprints202511.0268.v1

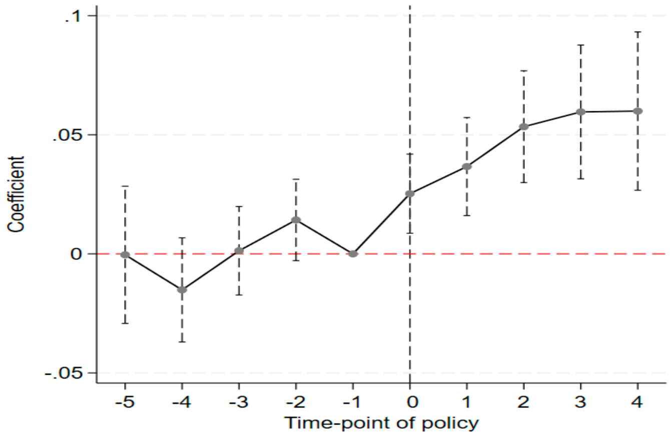

Figure 1. Parallel trend test.

4.2.2. Placebo Test

To further eliminate the potential influence of unobserved random variables on corporate low-carbon technological innovation, this study employs a placebo test by randomly assigning both the timing and location of public data releases. Specifically, we perform 500 random sampling iterations. In each iteration, 26 provinces are randomly selected and assigned fictitious launch dates to form a virtual treatment group, while the remaining 5 provinces constitute the virtual control group. The benchmark regression model is then re-estimated.

Figure 2 presents the kernel density distributions of the estimated coefficients and P-values for the explanatory variables in the placebo test. The kernel density of the estimated coefficients nears a normal distribution, centered at zero. Only a very small number of regression coefficients exceed the actual regression coefficients listed in Column (2) of Table 2. Furthermore, the estimated coefficients are non-significant in most cases. This outcome reveals that the influence of open public data on low-carbon technological innovation is not coincidental or attributable to other unknown factors, thereby confirming the reliability of this study's findings.

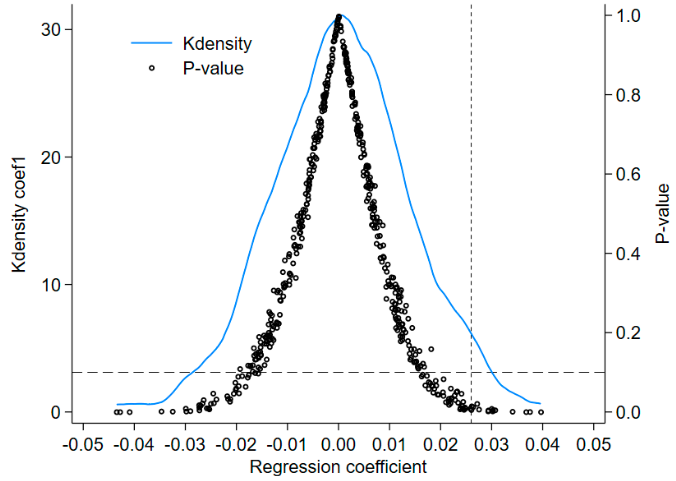

Figure 2. Placebo test.

4.2.2. Eliminate Other Policies Interference

To test whether the benchmark results are affected by other relevant policies, this study controls for several concurrent initiatives in columns (1) to (4) of Table 3: the Public Information Resource Opening Pilot Policy ($Pinfor$), the National Big Data Comprehensive Pilot Zone Policy ($Ndata$), the Smart City Pilot Policy ($Wcity$), and the Low-carbon City Pilot Policy ($Lcarcon$). Columns (1) to (4) in Table 3 display the regression results with each policy controlled individually, while column (5) presents the results controlling for all policies simultaneously. The results show that the estimated coefficients for $Open$ are all significantly positive, further validating the robustness of the benchmark regression results.

Table 3. Test outcomes isolating influence from other policies.

Variables	Lct				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Open	0.023***	0.026***	0.029***	0.031***	0.028***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.010)
Pinfor	-0.022				-0.028*
	(0.015)				(0.016)
Ndata		-0.007			-0.009
		(0.016)			(0.017)
Wcity			-0.014		-0.007
			(0.017)		(0.017)
Lcarbon				-0.023*	-0.024*
				(0.013)	(0.014)
Control variables	YES	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
Observations	37705	37705	34150	34150	34150
R-squared	0.638	0.638	0.642	0.642	0.642

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, Standard errors in parentheses are clustered at the firm level.

4.3. Other Robustness Tests

To further validate the robustness of the benchmark results, we conduct the following additional tests.

4.3.1. Two-Way Clustered Standard Errors

To account for potential intertemporal autocorrelation and intra-provincial correlation, we employ two-way clustered standard errors across firm and provincial dimensions in the benchmark regression. Columns (1) and (2) of Table 4 shows that the coefficients of $Open$ are statistically significant and positive at the 5% level, suggesting that variations in the clustering level of standard errors do not materially influence the conclusions of the baseline regression.

4.3.2. Adjust the Sample Period

Since Shanghai launched China's first provincial-level open public data platform in 2012, when very few platforms existed, this paper excludes samples prior to 2012 from the regression analysis. Column (3) of Table 4 presents the estimation results, indicating that the coefficient of $Open$ is 0.027 and achieves significance at the 1% level, further confirming the robustness of the benchmark regression conclusions.

4.3.1. PSM

Given that sample characteristics may introduce selection bias due to group incompatibility, this study employs the PSM–DID method to mitigate this bias. In particular, we use control variables as matching covariates and implement annual 1:1 nearest-neighbor matching. Subsequently, regression is performed on the matched observations that meet these requirements. The results presented in

 doi:10.20944/preprints202511.0268.v1

columns (4) of Table 4 are consistent with benchmark regression, reaffirming the robustness of the baseline results.

Table 4. Results of robustness tests.

Variables	Two-way clustered standard errors		Remove samples prior to 2012	PSM
	(1)	(2)	(3)	(4)
	Lct	Lct	Lct	Lct
Open	0.026**	0.026**	0.027***	0.027***
	(0.009)	(0.009)	(0.010)	(0.009)
Control variables	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Observations	37705	37705	32019	36336
R-squared	0.638	0.638	0.674	0.641

Notes: * p < 0.1, ** p < 0.05, *** p < 0.01, Standard errors in parentheses are clustered at the firm level.

4.4. Endogeneity Test

The above analyses control for multiple factors that may influence the baseline model's results across multiple dimensions, thereby alleviating endogeneity to some extent. However, potential biases arising from reverse causality may still exist. For example, leading low-carbon technology firms may have incentives to promote data openness through policy advocacy and other channels to strengthen their competitive advantage. Concurrently, governments may proactively release data to reinforce the success of their transformation efforts. This two-way dynamic creates a bidirectional interaction between corporate demand for low-carbon innovation and government data supply. To address this, we employ an instrumental variables approach, estimated using two-stage least squares (2SLS).

Following Lyu et al. [10], this study selects the density of long-distance optical cables ($Lable$) in the province where the firm is located as an instrumental variable. The rationale for its use is as follows: First, the density of optical cables reflects the level of regional information infrastructure and is directly related to the government's capacity to establish open public data platforms, thereby satisfying the relevance condition. Second, the construction of long-distance optical fiber cables in China is centrally planned and managed by the country's four major network operators. As a result, individual firms have limited influence over this process, and the density of long-distance optical cables is therefore unlikely to directly affect enterprise low-carbon technological innovation, thereby fulfilling the exogeneity requirement.

Column (1) of Table 5 presents the regression results for the first stage. At the 1% significance level, the instrumental variable shows a significant positive correlation with public data openness, consistent with theoretical expectations. Column (2) of Table 5 presents the regression results for the second stage. The coefficient of $Open$ is significantly positive at the 5% level, consistent with the direction observed in the benchmark regression. Furthermore, the Kleibergen-Paap rk LM statistic is significantly positive at the 1% level, and the Cragg-Donald Wald F-statistic is 333.67, exceeding the threshold of 10. This eliminates concerns of under-identification and weak instrumental variables. These results indicate that, after addressing endogeneity, the main conclusions of this study remain valid.

 doi:10.20944/preprints202511.0268.v1

Table 5. Endogenous treatment estimation results.

Variables	First-Stage	Second-Stage
	(1)	(2)
	Open	Lct
Open		0.167** (0.0679)
Iv	0.893*** (0.0599)	
Control variables	YES	YES
Firm FE	YES	YES
Year FE	YES	YES
Kleibergen–Paap rk LM statistics		122.605***
Cragg-Donald Wald F statistic		442.057
Observations	37,705	37,705

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, Standard errors in parentheses are clustered at the firm level.

4.5. Mechanism Analysis

4.5.1. Government Transparency

In line with Hypothesis 2, if open public data enhances government transparency, then its effect on promoting corporate low-carbon technological innovation should be more pronounced in regions with lower initial transparency. To test this mechanism, this study uses the Government Transparency Index published by the Chinese Academy of Social Sciences to measure government transparency (Gt), with lower values indicating weaker transparency. Specifically, the variable Gt and its interaction term $Gt \times Open$ are introduced into Model (1) for estimation. As shown in Column (1) of Table 6, the coefficient of the interaction term is significantly negative, indicating that the positive effect of open public data on corporate low-carbon technological innovation is more pronounced when government transparency is lower. This finding partially supports Hypothesis 2 that open public data can foster corporate low-carbon technological innovation by enhancing government transparency.

4.5.2. Barriers to Factor Mobility

Consistent with Hypothesis 2, open public data is also expected to mitigate barriers to factor mobility. Referencing to Bresnahan and Reiss [28], we adopt the Herfindahl-Hirschman Index (HHI) as a proxy for barriers to factor mobility (Fmb). A higher HHI value indicates greater market concentration and, consequently, stronger barriers to factor mobility. Specifically, both Fmb and the interaction term $Fmb \times Open$ are incorporated into Model (1) for estimation. As shown in column (2) of Table 6, the coefficient of the interaction term is significantly positive, implying that the promotional effect of open public data on corporate low-carbon technological innovation is stronger in regions with higher market concentration—that is, where barriers to factor mobility are greater. This finding partially supports Hypothesis 2 that open public data can foster corporate low-carbon technological innovation by reducing barriers to factor mobility.

 doi:10.20944/preprints202511.0268.v1

Table 6. Mechanism test results.

Variables	Lct	
	(1)	(2)
Open	0.039***	0.012
	(0.009)	(0.010)
Gt	0.022**	
	(0.009)	
Open*Gt	-0.027***	
	(0.010)	
Fmb		-0.025**
		(0.010)
Open*Fmb		0.031***
		(0.012)
Control variables	YES	YES
Firm FE	YES	YES
Year FE	YES	YES
Observations	29629	36715
R-squared	0.674	0.638

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, Standard errors in parentheses are clustered at the firm level.

4.6. Heterogeneity Analysis

Low-carbon technological innovation depends not only on access to external data but also on firms' internal resources and capabilities. First, as low-carbon technological innovation involves long cycles, high risks, and considerable uncertainty [29], R&D activities demand substantial financial resources. Firms with greater R&D investment are more likely to engage in sustained innovation activities [30]. Second, the limited capacity of internal financing forces firms to rely on external capital to pursue high-quality innovation [31]. An improved ability to obtain financing not only lowers the cost of innovation funding but also enables firms to scale up their financing, thereby providing more robust support for high-quality innovation activities. Finally, by leveraging digital technologies, firms can obtain critical information on changes in the external environment, market dynamics, and technological advancements, which in turn helps lower sunk costs and reduce the risks associated with innovation [32]. Accordingly, this study employs three firm-level variables—R&D investment, financing constraints, and digitalization level—to capture the heterogeneous effects of firm capabilities. This analysis elucidates how firms integrate, build, and reconfigure internal and external resources to adapt to rapidly changing environments.

4.6.1. R&D Investment

The logarithmic value of R&D expenditure is used to measure firms' R&D investment. Specifically, the sample is divided into high- and low-R&D investment groups based on whether a firm's R&D investment exceeds the sample median. The results are presented in columns (1) and (2) of Table 7. The estimated coefficient of $Open$ is significantly positive across both subsamples. However, the Chow test confirms a statistically significant difference between them. The results show that the effect of open public data is greater for firms with high R&D investment than for those with low R&D investment.

4.6.2. Financing Constraints

Following Kaplan and Zingales [33], we use the KZ index to measure firms' financing constraints, with a lower value indicating weaker constraints. Specifically, firms are divided into high- and low-constraint groups based on whether their KZ index exceeds the sample median. The results are presented in columns (3) and (4) of Table 7. The estimated coefficient of $Open$ is

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 November 2025 doi:10.20944/preprints202511.0268.v1

significantly positive only for firms facing lower financing constraints. This suggests that, relative to firms with higher financing constraints, open public data primarily promotes low-carbon technological innovation among those with weaker financing constraints.

4.6.3. Digitalization Level

Following He et al. [34], this study develops a comprehensive index system to assess digital transformation by employing a text analysis approach. Specifically, the sample is divided into high- and low-digitalization groups based on whether a firm's digital transformation index exceeds the sample median. Columns (5) and (6) of Table 7 indicate that the estimated coefficient of $Open$ is significantly positive only for firms with higher levels of digitalization. This indicates that, relative to firms with lower digitalization levels, open public data primarily promotes low-carbon technological innovation among firms with higher levels of digitalization, underscoring the amplifying role of digital capability.

Table 7. Heterogeneity analysis results.

Variables	R&D investment		Financing Constraints		digitalization	
	Low	High	Low	High	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)
Open	0.022**	0.042**	0.037***	0.007	0.017	0.022*
	(0.009)	(0.017)	(0.012)	(0.013)	(0.011)	(0.012)
Control variables	YES	YES	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Observations	13654	13777	18426	18419	16866	15995
R-squared	0.572	0.709	0.666	0.671	0.610	0.728
p-value of Chow test	0.000		0.029		0.061	

Notes: * p < 0.1, ** p < 0.05, *** p < 0.01, Standard errors in parentheses are clustered at the firm level.

5. Further Discussion: Non-Rival Spillover Effects of Public Data Openness

The non-rivalrous nature of public data allows its value to extend beyond direct beneficiaries, creating the potential for extensive spillover effects through knowledge diffusion. Accordingly, this section extends the analysis along two dimensions: geographic and industrial. First, it investigates whether open public data can transcend administrative boundaries to foster regional collaborative innovation. Second, it explores whether data openness induces technological imitation and competitive convergence within industries, thereby reshaping the broader innovation ecosystem.

5.1. Geographical Spillover Effect

Following Alder et al. [35], this study constructs the following model to analyze the spatial spillover effects of public data openness:

$$Lct_{it} = \beta_0 + \beta_1 Open_{it} + \sum_{s=200}^{1200} \delta_s \times N_{it}^s + \beta_2 Control_{it} + \lambda_i + \mu_t + \varepsilon_{it} \qquad (4)$$

Where s denotes the geographical distance between provinces (in kilometers, $s \geq 200$), measured by the straight-line distance between any two provincial capitals. δ_s reflects the impact of public data openness on low-carbon technological innovation among firms in other regions. N_{it}^s is the geographic spillover variable, taking a value of 1 if at least one provincial public data openness platform exists within s kilometers of firm i's location (excluding its own province), and 0 otherwise (s=200, 400, ⋯, 1000). The remaining variables are defined consistent with model (1).

As shown in Figure 3, the geographic spillover effects of public data openness exhibit an inverted U-shaped pattern. First, there exists a localized promotion effect: within the 400 – 600 km range, the coefficient δ_s is 0.028 and statistically significant at the 1% level. This occurs because a medium distance achieves a balance between the knowledge diffusion effect and the competitive suppression effect—proximity may lead to resource siphoning, whereas excessive distance results in information distortion. Second, a boundary attenuation effect is observed. When the distance is below 400 km or beyond 600 km, the coefficients become insignificant, confirming the existence of a "critical spillover radius" in geoeconomic space. As distance increases, spillover effects display a nonlinear pattern of "insignificant → significant → insignificant," rather than a monotonic decline.

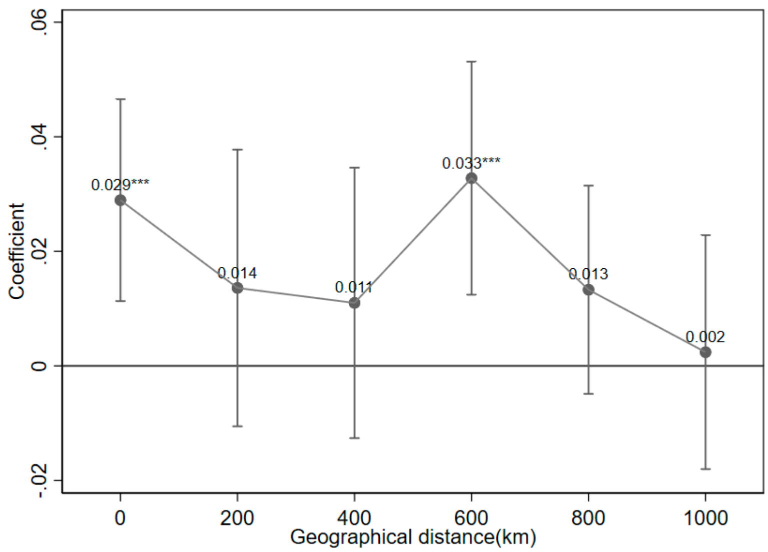

Figure 3. Geographical Spillover Effect.

5.2. Industry Spillover Effect

To examine the industrial spillover effects of public data openness, this study follows Li et al. [36] and constructs the following model:

$$Lct_{it} = \beta_0 + \beta_1 Open_{it} + \beta_2 Peer_{it} + \beta_3 Control_{it} + \lambda_i + \mu_{pt} + \varepsilon_{it} \quad (5)$$

$$Peer_{it} = \frac{n\_treatfirms_{mt} - Open_{it}}{n\_firms_{mt} - 1} \quad (6)$$

where m denotes industry. $Peer_{it}$ represents the extent to which firms in the same industry as firm i are affected by the launch of provincial public data platforms in year t, thereby capturing the industry peer effect of public data openness. Specifically, n_firms_{mt} denotes the total number of firms in industry m during year t; $n\_treatfirms_{mt}$ represents the number of firms in industry m during year t that are affected by the launch of the provincial public data platform in their location, and $(n\_treatfirms_{mt} - Open_{it})$ can be interpreted as the number of other firms–excluding firm i—in industry m and year t that are influenced by the launch of provincial public data platforms in their respective provinces. The remaining variables are defined consistently with model (1).

As reported in Table 8, the regression results show that the coefficients for both $Open$ and $Peer$ are significantly positive, demonstrating that the impact of public data platforms exhibits a significant within-industry peer effect. Specifically, as more peer firms within the same industry are affected by the establishment of provincial-level public data platforms, individual firms exhibit greater low-carbon technological innovation. Overall, these results suggest that the non-rival nature of public data reshapes industrial innovation ecosystems through mechanisms of technological imitation and knowledge diffusion.

Table 8. Industry Spillover Effect.

Variables	Lct	
	(1)	(2)
Open	0.025***	0.025***
	(0.009)	(0.009)
Peer	0.065**	0.054*
	(0.031)	(0.031)
Control variables	YES	YES
Firm FE	YES	YES
Year FE	YES	YES
Industry FE	YES	YES
Observations	37600	37600
R-squared	0.639	0.641

Notes: * p < 0.1, ** p < 0.05, *** p < 0.01, Standard errors in parentheses are clustered at the firm level.

6. Conclusions and Policy Implications

6.1. Conclusions

This study leverages China's local government open public data platforms as a quasi-natural experiment to systematically examine the impact of public data openness on corporate low-carbon technological innovation and its underlying mechanisms. The main research conclusions are as follows.

First, public data openness significantly promotes corporate low-carbon technological innovation, and this finding remains robust after a series of validation and robustness checks. Second, mechanism analysis reveals that public data openness promotes corporate low-carbon technological innovation by enhancing government transparency and reducing barriers to factor mobility. Third, the impact of public data openness policies is highly heterogeneous, exerting a more pronounced effect on low-carbon technological innovation in firms characterized by higher R&D investment, greater digitalization, and fewer financing constraints. Finally, the positive impact of public data openness on low-carbon technological innovation exhibits significant spillover effects across both geographic and industrial dimensions. Geographically, the optimal spillover radius lies between 400 and 600 kilometers. Industrially, greater policy coverage of data openness within the same sector significantly stimulates low-carbon innovation even among firms not directly subject to such policies.

6.2. Policy Implications

Based on the above conclusions, the following policy recommendations are proposed. First, Policymakers should deepen data openness to construct a high-quality supply system. Although the pilot policy has been steadily advancing, many provinces and municipalities have not yet participated in the construction of public data openness platforms. Therefore, the government should establish and continuously expand a nationwide, standardized National Low-Carbon Innovation Public Data Resource Repository. This repository should prioritize integration of high-value datasets essential to corporate low-carbon innovation, including energy consumption and carbon emissions monitoring, as well as datasets on green supply chains. It must ensure the data's authority, timeliness, and machine-readability to reduce institutional barriers to corporate innovation at the source.

Second, the government should focus on strengthening firms' ability to convert data elements into innovation value. For high-capability firms—those with strong R&D investment and advanced digitalization—governments should incentivize them to lead in forming industrial data innovation consortia, thereby sharing data dividends. For financially constrained firms, policymakers should design "data-finance linkage" instruments. These could include incorporating public-data-driven green technology projects into green credit subsidies and risk compensation schemes, thus mitigating the adverse impact of financing constraints on policy effectiveness.

 doi:10.20944/preprints202511.0268.v1

Finally, to maximize the systemic value of data openness, governments should foster cooperation across regions and industries. The study finds that the policy impact has an optimal spillover radius of 400–600 kilometers and exhibits significant network effects across industries. Therefore, policymakers should adopt a dual-coordination strategy. Geographically, efforts should focus on guiding the efficient allocation of data resources within the optimal spillover range. At the industrial level, governments should support high-carbon sectors in building low-carbon data value chains that drive precise emission reduction and collaborative innovation along the industry chain. By fostering an integrated network that connects geographic and industrial dimensions, the compound value and systemic effectiveness of public data openness can be maximized.

6.3. Limitations

This study has several limitations. First, although the study verifies the positive effects of open public data on corporate low-carbon technological innovation, it does not distinguish the heterogeneous impacts of different types of data elements. Second, the analysis focuses on whether data are open, without further exploring the role of data openness quality. Future research could extend this work by incorporating cross-country samples, evaluating the classification of data elements, and examining the mechanisms through which data quality influences low-carbon innovation.

Author Contributions: Conceptualization, Jing Wang and Z.C.; methodology, Jing Wang, Jie Wang and Z.C.; formal analysis, Jing Wang and Z.C.; resources, Jing Wang, Jie Wang and Z.C.; writing—original draft preparation, Jing Wang; writing—review and editing, Z.C. and Jie Wang; supervision, Z.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by (1) Anhui Provincial Philosophy and Social Sciences Planning Project, grant number AHSKQ2024D164. (2) National Social Science Fund, grant number 25CJY156.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data supporting reported results are available upon request from the authors.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Wang, H.; Feng, T.; Kong, J.; Cui, M.; Xu, M. Grappling with the trade-offs of carbon emission trading and green certificate: Achieving carbon neutrality in China. *J. Environ. Manag.* **2024**, *360*, 121101.
2. Malhotra, A.; Schmidt, T.S. Accelerating low-carbon innovation. *Joule* **2020**, *4*, 2259-2267.
3. Mao, Y.; Lin, Y. Do more hands make work easier? Public supervision and corporate green innovation. *Int. Rev. Econ. Finance* **2024**, *91*, 1064-1083.
4. Jones, C.I.; Tonetti, C. Nonrivalry and the economics of data. *Am. Econ. Rev.* **2020**, *110*, 2819-2858.
5. Zuiderwijk, A.; Janssen, M. Open data policies, their implementation and impact: A framework for comparison. *Gov. Inf. Q.* **2014**, *31*, 17-29.
6. Nagaraj, A. The Private Impact of Public Data: Landsat Satellite Maps Increased Gold Discoveries and Encouraged Entry. *Manage. Sci.* **2022**, *68*, 564-582.
7. Magalhaes, G.; Roseira, C. Open government data and the private sector: An empirical view on business models and value creation. *Gov. Inf. Q.* **2020**, *37*, 101248.
8. Yan, R.; Xing, C.; Chen, X.; Zhao, Y. Is it real or illusory? An empirical examination of the impact of open government data on innovation capability in the case of China. *Technol. Soc.* **2023**, *75*, 102396.

 doi:10.20944/preprints202511.0268.v1

9. Chen, Z.; Zhang, Y.; Wang, H.; Ouyang, X.; Xie, Y. Can green credit policy promote low-carbon technology innovation? *J. Clean. Prod.* **2022**, *359*, 132061.
10. Lyu, Y.; Bai, Y.; Zhang, J. Digital transformation and enterprise low-carbon innovation: A new perspective from innovation motivation. *J. Environ. Manag.* **2024**, *365*, 121663.
11. Yang, G.; Nie, Y.; Li, H.; Wang, H. Digital transformation and low-carbon technology innovation in manufacturing firms: The mediating role of dynamic capabilities. *Int. J. Prod. Econ.* **2023**, *263*, 108969.
12. Wang, M.; Li, Y.; Li, M.; Shi, W.; Quan, S. Will carbon tax affect the strategy and performance of Low-carbon technology sharing between Enterprises? *J. Clean. Prod.* **2019**, *210*, 724–737.
13. Luo, Y.; Liu, Y.; Wang, D.; Han, W. Low-carbon city pilot policy and enterprise low-carbon innovation–A quasi-natural experiment from China. *Econ. Anal. Policy* **2024**, *83*, 204–222.
14. Qi, S.Z.; Zhou, C.B.; Li, K.; Tang, S.Y. Influence of a pilot carbon trading policy on enterprises' low-carbon innovation in China. *Clim. Policy* **2021**, *21*, 318-336.
15. Hughes-Cromwick, E.; Coronado, J. The value of US government data to US business decisions. *J. Econ. Perspect.* **2019**, *33*, 131-146.
16. Schmidthuber, L.; Ingrams, A.; Hilgers, D. Government openness and public trust: The mediating role of democratic capacity. *Public Adm. Rev.* **2021**, *81*, 91-109.
17. Goldfarb, A.; Tucker, C. Digital economics. *J. Econ. Lit.* **2019**, *57*, 3-43.
18. Shane, S. Prior knowledge and the discovery of entrepreneurial opportunities. *Organ. Sci.* **2000**, *11*, 448-469.
19. Jetzek, T.; Avital, M.; Bjorn-Andersen, N. Data-driven innovation through open government data. *J. Theor. Appl. Electron. Commer. Res.* **2014**, *9*, 100-120.
20. Yu, Y.; Xie, B.; Dou, Z.; Fu, Q. Managerial myopia and corporate innovation strategy. *Finance Res. Lett.* **2024**, *67*, 105733.
21. Park, S.; Gil-Garcia, J.R. Open data innovation: Visualizations and process redesign as a way to bridge the transparency-accountability gap. *Gov. Inf. Q.* **2022**, *39*, 101456.
22. Akerlof, G. The Market for "Lemons": Quality Uncertainty and the Market Mechanism. *Q. J. Econ.* **1970**, *84*, 488-500.
23. Farboodi, M.; Veldkamp, L. A model of the data economy. National Bureau of Economic Research: Cambridge, MA, USA, 2021.
24. Teece, D.J.; Pisano, G.; Shuen, A. Dynamic capabilities and strategic management. *Strateg. Manage. J.* **1997**, *18*, 509-533.
25. Nikiforova, A.; Lnenicka, M. A multi-perspective knowledge-driven approach for analysis of the demand side of the Open Government Data portal. *Gov. Inf. Q.* **2021**, *38*, 101622.
26. Zhu, J.; Fan, Y.; Deng, X.; Xue, L. Low-carbon innovation induced by emissions trading in China. *Nat. Commun.* **2019**, *10*, 4088.
27. Beck, T.; Levine, R.; Levkov, A. Big bad banks? The winners and losers from bank deregulation in the United States. *J. Finance* **2010**, *65*, 1637-1667.
28. Bresnahan, T.F.; Reiss, P.C. Entry and competition in concentrated markets. *J. Polit. Econ.* **1991**, *99*, 977-1009.
29. Weber, T.A.; Neuhoff, K. Carbon markets and technological innovation. *J. Environ. Econ. Manage.* **2010**, *60*, 115-132.
30. Wei, J.; Li, Y.; Liu, X.; Du, Y. Enterprise characteristics and external influencing factors of sustainable innovation: based on China's innovation survey. *J. Clean. Prod.* **2022**, *372*, 133461.
31. Bhattacharya, U.; Hsu, P.H.; Tian, X.; Xu, Y. What affects innovation more: policy or policy uncertainty? *J. Financ. Quant. Anal.* **2017**, *52*, 1869-1901.

 doi:10.20944/preprints202511.0268.v1

32. Bharadwaj, A.; El Sawy, O.A.; Pavlou, P.A.; Venkatraman, N.V. Digital business strategy: toward a next generation of insights. *MIS Q*. **2013**, 471-482.
33. Kaplan, S.N.; Zingales, L. Do investment-cash flow sensitivities provide useful measures of financing constraints? *Q. J. Econ.* **1997**, *112*, 169-215.
34. He, G.; Li, Z.; Yu, L.; Zhou, Z. Does commercial reform embracing digital technologies mitigate stock price crash risk? *J. Corp. Finance* **2025**, *91*, 102741.
35. Alder, S.; Shao, L.; Zilibotti, F. Economic reforms and industrial policy in a panel of Chinese cities. *J. Econ. Growth* **2016**, *21*, 305-349.
36. Li, X.; Yuan, C.; Zhao, Y.; Wu, M.; Cao, A.; Liu, L. Peer effects in urban pollution and carbon reduction: evidence from China. *Sustain. Cities Soc.* **2025**, 106521.