Inzinerine Ekonomika-Engineering Economics, 2026, 37(1), 106–122

Data-Driven Empowerment: How Government Data Governance Facilitates Firm Upgrading

Dan Wang[1], Jian Hu[2], Shenglin Ma[3*]

[1]*School of Economics, Guangxi University*
100 Daxue East Road, Xixiangtang District, Nanning, Guangxi, 530004, P. R. China
E-mail. 2330401007@st.gxu.edu.cn

[2]*School of Statistics, Tianjin University of Finance and Economics*
25 Zhujiang Road, Hexi District, Tianjin, 300222, P. R. China
E-mail. xahja09@stu.tjufe.edu.cn

[3]*School of Economics and Management, North University of China*
Xueyuan Road, Jiancaoping District, Taiyuan, Shanxi, 030051, P. R. China
*E-mail. shenglinma5364@yeah.net (*corresponding author)*

https://doi.org/10.5755/j01.ee.37.1.41789

Promoting deep industrial transformation and upgrading serves as the micro-level foundation for improving the institutional framework of new quality productive forces and advancing a high-standard socialist market economy. Enhancing government data governance mechanisms represents a critical pathway to achieving this objective. Drawing on panel data from 2011 to 2021 concerning the establishment of government big data administration agencies and Chinese A-share listed firms, this paper employs a multi-period difference-in-differences (DID) approach to systematically examine the impact of government data governance on firm upgrading, along with its underlying mechanisms. The empirical results reveal that government data governance significantly facilitates firm upgrading, and this effect remains robust across a series of sensitivity checks. Heterogeneity analyses indicate that the positive effects are more pronounced among firms in heavily regulated industries, high-tech sectors, enterprises with higher levels of data utilization, and regions with higher administrative levels of data governance institutions. Mechanism tests further suggest that government data governance improves the quality of public service delivery and empowers corporate digital governance, both of which play key intermediary roles in promoting firm upgrading. These findings confirm the enabling role of government data governance in advancing firm upgrading and offer policy implications for stimulating innovation among market entities and strengthening the micro-foundations of China's economic recovery and growth momentum.

Keywords: *Government Data Governance; Firm Upgrading; Big Data Administration Agencies*.

Introduction

Enhancing government data governance mechanisms has emerged as a pivotal strategy to improve administrative efficiency and public service delivery, as well as to advance the modernization of national governance systems. China's 2025 Government Work Report explicitly calls for accelerating the establishment of foundational data institutions and promoting the development and utilization of data resources. On October 25, 2023, the official establishment of the National Data Administration marked a significant step toward building a functional data factor market and advancing the broader vision of "Digital China." In the same year, China's total data output reached 32.85 ZB, representing a 22.44 % year-on-year increase, underscoring the growing strategic value of data as a production factor in resource allocation and deep industrial transformation. These developments raise a central research question: can government data governance, beyond improving bureaucratic efficiency, serve as an institutional driver of enterprise upgrading and innovation?

Against this backdrop, firms, as the most dynamic microeconomic agents, play a critical role in driving structural upgrading and fostering new quality productive forces. Prior studies have highlighted that firm upgrading is a key mechanism through which economic restructuring and competitive advantage are realized (Liang & Lu 2024; Zhu & Tomasi 2020). From the perspective of information economics and transaction cost theory, government data governance may promote firm upgrading by reshaping the information environment and reducing institutional frictions. Moreover, it serves as an essential indicator of the evolution of new quality productive capacities. Despite the growing academic attention to firm upgrading, the literature has yet to reach a consensus on its conceptual definition. Many scholars interpret firm upgrading as a process of technological and organizational advancement. In this process, enterprises evolve from low-tech, low-value-added production to high-tech, high-value-added activities, ultimately achieving sustainable growth and contributing to high-quality economic development (Verhoogen, 2023; Zhou *et al.,* 2019; Zhu & Yu, 2024).

Existing scholarship increasingly explores the interplay between digital government and enterprise behavior. The first stream of research on digital government and public management innovation emphasizes how e-government initiatives improve administrative efficiency, information transparency, and policy responsiveness (Mergel, 2019; Ripamonti, 2024). These studies have provided valuable theoretical foundations for understanding how digital technologies transform bureaucratic systems, yet they primarily remain at the macro level and seldom capture firm-level dynamics. A second line of inquiry examines governance and firm behavior at the micro level. Scholars in this domain highlight that the institutional quality of public governance, regulatory transparency, and government–firm interactions significantly influence firm investment, innovation, and upgrading decisions (Li & Yue, 2025; Zhu & Yu, 2024). Drawing on information economics and transaction cost theory, this literature underscores that a well-functioning digital governance framework can mitigate information asymmetry, enhance coordination efficiency, and thereby create conditions conducive to enterprise upgrading. A third research stream focuses on the determinants of enterprise upgrading, emphasizing that innovation capability, efficient resource allocation, and a supportive institutional environment are central to technological progress and organizational transformation (Liang & Lu, 2024; Lo *et al.,* 2023; Luo *et al.,* 2023; Ma *et al.,* 2025; Verhoogen, 2023).

Despite these valuable contributions, an integrated perspective on how government data governance, situated at the intersection of public digitalization and corporate transformation, influences enterprise upgrading remains limited. Few studies have systematically explored how data governance mechanisms restructure information flows, reduce transaction costs, and enhance firms' digital capabilities, particularly within developing economies. This research gap highlights the need to investigate government data governance as an independent institutional pathway through which digital governance promotes firm upgrading and supports the development of new quality productive forces.

In China's transitional economy, the government has played a pivotal role in shaping firm upgrading trajectories. The state controls nearly 80 % of national data resources and wields substantial discretionary power in administrative approvals, the allocation of scarce production factors, and policy implementation. The quality of public service provision constructed by subnational governments has become a critical determinant of firms' upgrading decisions (Zhu & Yu, 2024). By the end of 2023, more than thirty provincial-level administrative regions had established dedicated data management agencies, demonstrating the institutionalization of data governance at scale. Against this institutional backdrop, a key and pressing question arises: How can the government, through public data governance, enable "data-based decision-making, data-driven management, and data-enabled innovation," thereby fostering a virtuous cycle of state–firm interaction that supports firm upgrading? Understanding this mechanism is essential for informing both theoretical development and practical policy design in the context of digital governance and economic modernization.

With the rapid advancement of digital technologies, government data governance has emerged as a new engine for transforming public governance paradigms and modernizing national governance systems. In 2015, the State Council issued the Action Plan for Promoting Big Data Development, which for the first time explicitly advocated the principle of "governing with data, deciding with data, managing with data, and innovating with data." Subsequently, the Guiding Opinions on Strengthening Digital Government Development, released in June 2022, further emphasized the need to "apply digital technologies widely in government management and services," "advance digital and intelligent governance," and "build an open and shared data resource system."

The common goal of these policy directives is to enhance government service capabilities through data governance and to digitally restructure administrative functions and operational models. Government data governance, in this regard, serves multiple functions: it promotes information transparency, facilitates innovative government–firm interactions, and improves policy predictability. Moreover, it fosters cross-level and cross-agency information sharing and collaborative governance, thereby enhancing the overall efficiency of public service delivery. Thus, government data governance is not merely an instrument for improving bureaucratic performance; it also serves as a critical lever for facilitating firm upgrading and achieving high-quality economic development.

This study exploits the staggered establishment of local government big data administration agencies as a quasi-natural experiment and adopts a multi-period difference-in-differences (DID) framework to evaluate the impact of government data governance on firm upgrading. The empirical results show that government data governance significantly promotes firm upgrading. This core finding remains robust after a series of validity checks, including parallel trends testing and alternative specifications. Subsample analyses further reveal substantial heterogeneity in the effect of government data governance on firm upgrading. Specifically, the positive impact is more pronounced among firms in heavily regulated industries, high-tech sectors, and those with a higher intensity of data factor utilization. From a regional perspective, the upgrading effect is significantly stronger in jurisdictions where data governance institutions hold higher administrative ranks. Mechanism analyses indicate that improvements in the public service environment and the empowerment of firms' digital governance capabilities are two key channels through which government data governance facilitates firm upgrading.

It has become increasingly critical to understand how digital government initiatives affect firm upgrading in the era of digital transformation. This is essential from a theoretical, practical, and policy standpoint, as these initiatives work by optimizing the public service environment and empowering corporate digital governance. Clarifying the underlying mechanisms through which government data governance shapes firm behavior and facilitates upgrading is both timely and essential. This study seeks to address three core questions: How does government data governance influence firm upgrading? What are the mechanisms underlying this effect? And how can firms leverage the institutional presence of big data administration agencies to accelerate transformation and achieve high-quality development? These questions serve as the foundation for the empirical investigation presented in this paper.

Compared with the existing literature, this study makes several marginal contributions:

First, from a research perspective, this study extends the literature on digital government by moving beyond macro-level administrative efficiency and policy performance to explore government data governance and firm upgrading. While prior studies have predominantly examined digital government from a macro-administrative perspective, this paper extends the discussion to the micro level of enterprise development, providing a new lens to understand how state governance capacity translates into corporate transformation outcomes. In doing so, this research contributes to the theoretical development of government data governance by conceptualizing it as an institutional mechanism that reshapes information structures, enhances transparency, and reduces transaction costs, thereby offering a deeper explanation of how digital governance facilitates firm upgrading.

Second, from a methodological perspective, this study employs a multi-period difference-in-differences (DID) framework exploiting the quasi-natural experiment of staggered big data administration agency establishment across Chinese cities. This approach, complemented by PSM-DID, instrumental variable estimation, and heterogeneity-robust estimators, enables credible causal identification while addressing endogeneity concerns that limit cross-sectional analyses.

Third, from a content perspective, this study constructs a comprehensive theoretical framework linking institutional transformation to firm behavior through two mechanisms: optimizing the public service environment and empowering firms' digital governance capabilities. In addition, this study enriches the literature on corporate transformation and upgrading by identifying how external digital governance capacity interacts with firms' internal digital capabilities to drive innovation and structural upgrading. This dual-mechanism framework advances the understanding of how government-led data governance influences enterprise transformation pathways, while the heterogeneity analysis reveals its differentiated effects across regulated industries, high-tech sectors, and administrative hierarchies. These findings contribute to the literature by uncovering the institutional logic and practical conditions under which data governance enhances firm upgrading, providing actionable implications for cultivating new quality productive forces and promoting high-quality development in the digital era.

The remainder of the paper proceeds as follows. Section 2 develops the theoretical framework and outlines the hypotheses. Section 3 describes the empirical model and data construction. Section 4 presents the baseline results, robustness checks, and endogeneity mitigation strategies. Section 5 provides the mechanism and heterogeneity analyses. Section 6 concludes by summarizing the main findings, offering policy recommendations.

Theoretical Framework and Research Hypotheses

Government data governance represents a transformation and optimization of governmental processes and administrative operations through digital technologies. As a key initiative for reshaping public governance paradigms and advancing the modernization of the national governance system, it signifies a shift toward a new model of digital governance. Broadly, government data governance promotes firm upgrading through two principal mechanisms: optimizing the public service environment, and empowering firms' digital governance capabilities.

These two mechanisms are not mutually exclusive but complementary. The optimization of the public service environment reduces external information asymmetries, while the empowerment of firms' digital governance enhances internal information processing efficiency, together forming an integrated framework that lowers transaction costs and strengthens firms' adaptive capabilities.

Optimizing the Public Service Environment

This mechanism draws on information economics and public management theory, emphasizing that transparent and reliable information reduces search, coordination, and compliance costs for firms. First, government data governance enhances information transparency and improves policy predictability. According to the Guidelines for the Development of a National Integrated Government Big Data System, a core objective of data governance is to build a unified and standardized system for public data openness, providing accessible and regulated data services to market participants. This approach contributes directly to improving government information disclosure. On the one hand, governments consolidate vast volumes of fragmented data across departments and administrative levels through integrated e-government platforms and inter-agency data exchange systems, thereby increasing the availability and accessibility of public information (Ma & Appolloni, 2025; Ripamonti, 2024). On the other hand, firms can now access policy updates, submit feedback, and receive services more efficiently and in real time, thanks to the continued expansion of "Internet + Government Services" platforms and official digital channels, which include mobile apps and social media. These developments significantly reduce information asymmetries (Zhu & Yu, 2024). From the perspective of information economics, the enhancement of information transparency effectively lowers search and coordination costs, enabling firms to make more efficient production and investment decisions.

By reducing information acquisition costs and improving the timeliness of policy responses, firms can better anticipate regulatory adjustments and allocate resources more effectively in the face of uncertainty. A good example is the "Jiangsu Smart Administration Platform," which integrates data from taxation, customs, and market supervision departments. Through this system, manufacturing enterprises can access real-time policy updates and industrial guidance, allowing them to adapt investment and innovation strategies more efficiently. This case highlights how data governance enables a transparent, responsive policy environment that supports firms' strategic upgrading.

Second, government data governance helps enhance public service capacity and technical support systems for intellectual property (IP) protection. The government has implemented unified intellectual property service platforms, such as the Patent Service System introduced in 2023, to facilitate comprehensive digitalization of patent-related processes, encompassing application, examination, and registration. This development has significantly reduced institutional transaction costs for firms, improved administrative efficiency, and enhanced access to public services (Zhu & Yu, 2024). On the other hand, with the widespread application of technologies such as big data, cloud computing, and artificial intelligence in regulatory enforcement, the identification, monitoring, and penalization of IP infringements have been continuously upgraded. These technological advances have markedly improved the precision and deterrence of regulatory mechanisms. A more effective IP protection system increases firms' expected returns on innovation (Abdin *et al.,* 2024; Grimaldi *et al.,* 2021), thereby strengthening innovation incentives and facilitating the transformation and upgrading of knowledge-intensive enterprises. For instance, the Shanghai Data Exchange has provided technology firms with access to standardized datasets and secure transaction channels, which help protect proprietary data while enabling the commercialization of data assets. This institutional innovation demonstrates how data governance reforms can simultaneously safeguard innovation outcomes and stimulate enterprise upgrading.

Third, in the process of advancing data governance, local governments have invested heavily in digital infrastructure, laying the groundwork for enterprise digital transformation. For instance, the "Cloud Guizhou" platform in Guizhou Province has not only enhanced the integration capacity of public data resources but also provided firms with access to big data application platforms, institutional guarantees for data security, and shared technical services. These developments have spurred regional enterprise digitalization and industrial upgrading (Mergel, 2019).

A well-developed digital infrastructure facilitates the formation of an efficient, stable, and transparent data flow system, enhancing the government's capacity for precise governance and the private sector's access to digital resources. Consequently, data governance fosters a mutually reinforcing ecosystem in which public infrastructure investment and firm-level innovation jointly accelerate industrial upgrading.

Empowering Firms' Digital Governance

This mechanism is grounded in capability theory and innovation economics, suggesting that improved access to digital resources and institutionalized data standards strengthen firms' learning, innovation, and decision-making abilities. The digital transformation of firms relies heavily on the effective access to and efficient utilization of data factors. Existing studies suggest that capabilities in digital technology application, information resource management, digital infrastructure, and the implementation of digital strategies are foundational to successful digital transformation. The establishment of government big data administration agencies not only fulfills basic functions such as data integration, public service delivery, and informatization, but also assumes broader responsibilities in strategic planning, institutional design, infrastructure development, and digital talent cultivation (Ali et al., 2018; Mergel, 2019; Zeng *et al.,* 2025), thereby facilitating firms' digital upgrading.

These agencies serve as "connective institutions" that bridge government databases with enterprise innovation systems. For instance, data governance frameworks developed in Tianjin and Chengdu provide open-access interfaces for enterprises to utilize industrial and environmental datasets, improving firms' analytical capabilities and enabling data-driven innovation management. In essence, government data governance reshapes information structures by standardizing data formats, improving accessibility, and promoting interoperability, which collectively reduce transaction frictions and enhance firms' capacity for real-time decision-making.

Specifically, the government has enhanced fiscal support and policy incentives for emerging technologies, including big data, cloud computing, and artificial intelligence, while advancing the development of digital infrastructure, such as 5G networks, industrial internet platforms, and data centers. These efforts have created a more stable and efficient data environment for firms. Conversely, governments foster standardized and unified data governance frameworks, incorporating consistent data formats and interfaces, thereby mitigating institutional frictions in data acquisition and application. Firms are thus able to more efficiently integrate multi-source, cross-platform data, improve their analytics capabilities, and enhance their data-driven decision-making, all of which contribute to firm upgrading.

In addition, government data governance improves firms' internal digital governance capabilities by setting unified standards for data quality, privacy, and security management. This standardization reduces compliance uncertainty, enabling firms to allocate more attention to technological innovation rather than administrative coordination.

Furthermore, government data governance promotes the use of data as a key production factor. Data has become a central driver of new quality productive forces and is fundamental to improving firm productivity and operational efficiency. First, the government promotes public data openness and sharing, providing firms with timely, high-quality data resources at reduced search and acquisition costs. Additionally, local governments have implemented various data governance policies to establish efficient and compliant mechanisms for data circulation and trading, both on and off platforms, thereby enhancing the allocation efficiency of data resources and improving firms' access to data in terms of quantity and quality (Park & Gil-Garcia, 2022). Second, emerging regional data-trading pilots, such as those in Hainan and Chongqing, demonstrate how structured data exchange can generate new market-driven incentives for digital upgrading by encouraging enterprises to transform data from an auxiliary asset into a core factor of production. Third, the creation of data exchange and collaboration platforms fosters data interaction and business coordination between firms and between firms and government entities, enabling the formation of a low-cost, high-frequency digital application environment (Li & Yue, 2025).

Through these interactions, a feedback loop emerges: data governance enhances data accessibility and reliability, which promotes digital innovation, while firms' growing demand for high-quality data further drives the refinement of governance systems.

Based on the above theoretical mechanisms, we propose the following hypothesis:

Hypothesis 1: Controlling for other factors, government data governance promotes firm upgrading by optimizing the public service environment and empowering firms' digital governance capabilities.

Research Design

Model Specification

To empirically test the hypothesis developed above, this study exploits the staggered establishment of municipal big data administration agencies as a quasi-natural experiment. Following prior research, we adopt a multi-period difference-in-differences (DID) approach to identify the causal effect of government data governance on firm upgrading (Liang & Lu, 2024; Zhu & Yu, 2024). The baseline empirical model is specified as follows:

$$Upgrade_{i,t} = \alpha + \beta_1 Time \times Treat + \varphi Fc_{i,t} + \rho Cc_{c,t} + \mu_i + \lambda_t + \varepsilon_{i,t} \quad (1)$$

In this model, Upgrade denotes the level of firm upgrading. The interaction term Time×Treat is a binary indicator equal to 1 if firm i is located in a city that has established a big data administration agency in year t, capturing the intensity of local government data governance. The subscripts i and t refer to the firm and year, respectively. Fc represents a set of firm-level control variables, including firm size (Size), leverage ratio (Lev), firm age (Age), return on total assets (Roa), operating cash flow (Cfo), board size (Board), proportion of independent directors (Indep), ownership concentration (Top1), CEO duality (Dual), and Big Four audit dummy (Big4). Cc denotes city-level control variables, including the logarithm of regional GDP (Lngdp), government size (Gov), degree of openness (Open), and the number of internet broadband subscribers (Internet). ε is the error term. To mitigate potential omitted variable bias, the model includes firm fixed effects and year fixed effects, accounting for unobservable time-invariant firm characteristics and macroeconomic shocks. Standard errors are clustered at the city–year level to address potential serial correlation and heteroskedasticity among firms within the same city and year.

Variable Description

Dependent Variable

The dependent variable is firm upgrading (Upgrade). A key objective of firm upgrading is to enhance productivity through technological innovation, market expansion, and managerial improvement. Therefore, the enhancement of new quality productive forces is widely regarded as a direct manifestation of successful upgrading. As emphasized by President Xi Jinping, "new quality productive forces are fundamentally characterized by advanced combinations of labor, means of labor, and objects of labor, and are defined by significant improvements in total factor productivity (TFP). Innovation is their core attribute, quality is the key requirement, and their essence lies in advanced productivity." In theoretical terms, firm upgrading is a dynamic process that encompasses technological renewal, strategic market repositioning, and organizational transformation, with innovation at its core. Following existing studies, this paper constructs a composite index of new quality productive forces using the entropy method to proxy for firm upgrading. The detailed construction of this variable is presented in Table 1.

Table 1

Firm Upgrading Indicators

Primary Indicator	Secondary Indicator	Tertiary Indicator	Definition and Description	
Elemental Upgrading	Human Capital	R&D Personnel Compensation Ratio	(R&D expenditure − wage compensation) / operating revenue	+
		R&D Personnel Proportion	Number of R&D personnel / total number of employees	+
		Share of Highly Educated Employees	Number of employees with bachelor's degree or above / total number of employees	+
	Physical Capital	Fixed Assets Ratio	Fixed assets / total assets	+
		Manufacturing Expenditure Ratio	(Cash paid for operating activities + depreciation of fixed and intangible assets + impairment provision − cash paid for purchases and services) / (Cash paid for operating activities + depreciation of fixed and intangible assets + impairment provision)	+
Technological Upgrading	R&D Investment	R&D Depreciation Ratio	(R&D expenditure − depreciation) / operating revenue	+
		R&D Rental Expenditure Ratio	(R&D expenditure − rental expenses) / operating revenue	+
		R&D Direct Investment Ratio	(R&D expenditure − direct investment) / operating revenue	+
	Knowledge Capital	Intangible Assets Ratio	Intangible assets / total assets	+
Efficiency Upgrading	Capital Utilization Efficiency	Capital Turnover Ratio	Operating revenue / average total assets	+
		Inverse Equity Multiplier	Total assets / shareholders' equity	

Key Independent Variable

The key independent variable is Time×Treat, which captures the degree of government data governance, measured by whether a city has established a big data administration agency. Specifically, Time is a binary variable equal to 1 in year t and onwards if a city officially launched its big data administration agency in year t–1, and 0 otherwise. Treat identifies whether the firm belongs to the treatment group or control group; it equals 1 if the firm is located in a city that established a big data administration agency during the sample period, and 0 otherwise (Cheng *et al.,* 2025).

Considering the typical time lag between the formal inauguration and actual operational effectiveness of a big data administration agency, combined with the fact that many agencies are officially launched in the second half of the year, we define the year subsequent to official establishment as the effective implementation year when constructing our Time variable. The interaction term Time×Treat therefore denotes a difference-in-differences estimator, and the coefficient β_1 captures the effect of government data governance on firm upgrading. A significantly positive β_1 would support Hypothesis 1, indicating that the establishment of big data administration agencies significantly enhances firm upgrading in the respective jurisdiction.

The choice to use the establishment of local big data administration agencies as a proxy for government data governance is grounded in both institutional function and practical implementation:

First, from a functional perspective, the State Council's Guiding Opinions on Strengthening Digital Government Development (Document No. 14) explicitly mandates government departments to "strengthen data management responsibilities" and "coordinate the management of government, public, and social data." The strategic role of big data agencies in driving digital government transformation is underscored by this. For instance, in provinces such as Zhejiang, Guangdong, and Guangxi, local policy documents assign these agencies with key mandates. Such responsibilities include promoting digital government reform, leading regional digital economy development, and coordinating digital infrastructure. Existing literature confirms that such agencies commonly oversee four core functions: data management, e-government services, digital economy coordination, and public service delivery. These functions collectively support the broader goal of digital transformation.

Second, from an operational perspective, big data administration agencies aim to integrate and share data resources by establishing unified data platforms, thereby eliminating "information silos" across departments and reducing redundant digital government infrastructure. These agencies play a critical role in unlocking the potential of government data for enterprise development. Through institutional innovation and policy guidance, they promote the integration of the real and digital economies and encourage firms to adopt digitalized public services, thereby driving demand-side improvements in data governance. Based on the above rationale, the use of local big data administration agency establishment as a proxy for government data governance is both theoretically grounded and empirically appropriate.

Control Variables

This study incorporates a range of control variables that may influence firm upgrading (Liang and Lu, 2024; Wang and Yang, 2025). Specifically, we include the following firm-level controls: firm size (Size), leverage ratio (Lev), operating cash flow (Cfo), return on assets (Roa), ownership concentration measured by the shareholding ratio of the largest shareholder (Top1), CEO duality (Dual), board size (Board), proportion of independent directors (Indep), Big Four audit dummy (Big4), and firm age (Age). In addition, Equation (1) controls for city-level variables (City vars), including the level of regional economic development (Lngdp), internet penetration (Web), degree of openness (Open), and government size (Gov). To address potential issues of heteroskedasticity and serial correlation, we use robust standard errors and cluster the standard errors at the city–year level throughout the regressions. Detailed definitions and descriptions of the main variables used in Equation (1) are provided in Table 2.

Table 2

Variable Definitions and Descriptions

Variable	Definition
Panel A. Firm-Level Variables	
Upgrade	Firm upgrading, measured using a composite index constructed via the entropy method
Size	Firm size, measured as the natural logarithm of total assets at the end of the current period.
Lev	Leverage ratio, measured as total liabilities divided by total assets.
Age	Firm age, measured as the natural logarithm of the number of years since the firm's listing.
Roa	Return on assets, calculated as net profit divided by total assets.
Cfo	Operating cash flow, measured as net cash flow from operating activities divided by total assets.
Indep	Board independence, measured as the ratio of independent directors to total board members.
Board	Board size, measured as the natural logarithm of the number of board members.
Topsha	Ownership concentration, measured as the shareholding ratio of the largest shareholder.
Dual	CEO-chair duality, equal to 1 if the chairman and the CEO are the same person, and 0 otherwise.
Big4	Big Four auditor dummy, equal to 1 if the firm is audited by one of the Big Four accounting firms, and 0 otherwise.
Panel B. City-Level Variables	
Time×Treat	Difference-in-differences estimator; equal to 1 if city c established a big data administration agency in year t, and the observation is from year t+1 onward, 0 otherwise.
Lngdp	Regional economic development level, measured as the natural logarithm of regional GDP per capita.
Gov	Government size, measured as fiscal general public expenditure divided by regional GDP.
Open	Degree of openness, measured as actual foreign investment divided by regional GDP.
Internet	Level of internet development, measured as the number of broadband internet users per capita.

Sample Selection and Data Sources

Given data availability, this study selects all A-share listed firms on the Shanghai and Shenzhen Stock Exchanges from 2011 to 2021 as the research sample. Firm-level data are obtained from the China Stock Market and Accounting Research (CSMAR) database. City-level control variables are sourced from various issues of the China City Statistical

Yearbook. Information on the establishment of big data administration agencies is manually collected by the authors from official government websites, news reports, Baidu Encyclopedia, and other publicly accessible sources.

To ensure sample validity, the following screening criteria are applied to the initial dataset: (1) Firms in the financial industry are excluded; (2) Firms designated as ST, *ST, or PT (i.e., under special treatment for financial distress or irregular operations) are excluded; (3) Firms that simultaneously issue B-shares or H-shares are excluded; (4) Firms that have been listed for less than one year, delisted, or suspended from trading are excluded; (5) Firms with missing values for key variables are excluded. In addition, to mitigate the influence of extreme values, all continuous firm-level variables are winsorized at the 1% level. The final sample comprises 23797 firm-year observations.

Descriptive Statistics

Descriptive statistics for the main variables are presented in Table 3. The variable Upgrade ranges from 0.6138 to 14.4850, with a standard deviation of 2.4990, indicating substantial variation in the level of transformation and upgrading across firms. The mean of the key independent variable Time×Treat is 0.4477, suggesting that approximately 44.77 % of the sampled firms are located in cities that have established big data administration agencies. The coefficient of variation for Roa (return on assets) is 1.7562, reflecting notable differences in profitability among firms in the sample.

In terms of corporate governance structure: The average value of Indep is 0.3824, indicating that, on average, independent directors account for approximately one-third of board members. The mean of Topsha is 0.3401, suggesting that the largest shareholder holds, on average, 34.01 % of the company’s shares. Dual has a mean of 0.2831, meaning that in 28.31 % of the firms, the chairman and CEO positions are held by the same individual. The average value of Big4 is 0.0381, indicating that 3.81 % of the sampled firms are audited by one of the Big Four accounting firms. At the city level, the variable Internet, which measures the degree of internet development, ranges from 0.2417 to 18.9865. This wide span highlight significant disparities in digital infrastructure across cities.

Table 3

Descriptive Statistics of Key Variables

Variable	N	Mean	Median	Std. Dev.	Min	Max
Upgrade	23797	5.054	4.721	2.499	0.614	14.485
Time×Treat	23797	0.448	0.000	0.497	0.000	1.000
Size	23797	22.137	21.994	1.230	19.151	26.210
Lev	23797	0.423	0.413	0.208	0.051	0.932
Age	23797	2.155	2.303	0.791	0.000	3.332
Roa	23797	0.037	0.037	0.064	-0.258	0.218
Cfo	23797	0.046	0.045	0.069	-0.174	0.257
Indep	23797	0.382	0.364	0.073	0.250	0.600
Board	23797	2.284	2.303	0.252	1.609	2.890
Topsha	23797	0.340	0.317	0.149	0.084	0.750
Dual	23797	0.283	0.000	0.451	0.000	1.000
Big4	23797	0.038	0.000	0.191	0.000	1.000
Lngdp	23797	11.450	11.514	0.516	8.773	13.056
Gov	23797	0.158	0.143	0.066	0.044	2.223
Open	23797	0.026	0.023	0.018	0.000	0.229
Internet	23797	5.039	4.426	3.339	0.242	18.986

Empirical Results and Analysis

Baseline Regression Results

Table 4 reports the baseline regression results. Column (1) presents the estimation without any control variables. Column (2) includes firm-level controls, and the coefficient on the key independent variable Time×Treat remains significantly positive at the 5 % level. Column (3) further adds city-level control variables, and the coefficient on Time×Treat continues to be significantly positive at the 5 % level. Across all model specifications, the coefficient of the digital governance variable (Time×Treat) remains consistently positive and statistically significant at the 5 % level, suggesting that the establishment of big data administration agencies significantly promotes firm upgrading. These findings provide preliminary empirical support for Hypothesis 1.

From an economic perspective, the coefficient of 0.082 in Column (3) implies that, on average, firms located in cities with big data administration agencies have approximately 8.2 % higher upgrading levels compared with those in cities without such institutions. In addition, variance inflation factor (VIF) tests are conducted to assess potential multicollinearity among explanatory variables. The maximum VIF is only 2.30, and the mean VIF across all variables is 1.36, well below the commonly accepted threshold of 10. This suggests that multicollinearity is not a serious concern in the baseline model.

Table 4

Baseline Results

Variable	(1)	(2)	(3)
	Upgrade	Upgrade	Upgrade
Time×Treat	0.974***	0.080**	0.082**
	(0.032)	(0.040)	(0.040)
Size		0.118***	0.116***
		(0.036)	(0.036)
Lev		0.090	0.078
		(0.137)	(0.138)
Age		0.785***	0.785***
		(0.060)	(0.060)
Roa		-2.142***	-2.136***
		(0.255)	(0.254)
Cfo		1.140***	1.146***
		(0.195)	(0.193)
Indep		0.366**	0.362**
		(0.163)	(0.163)
Board		0.020	0.018
		(0.058)	(0.058)
Topsha		0.088	0.105
		(0.219)	(0.220)
Dual		-0.065**	-0.065**
		(0.032)	(0.032)
Big4		-0.056	-0.063
		(0.103)	(0.103)
Lngdp			0.138
			(0.118)
Gov			0.109
			(0.221)
Open			-1.088
			(1.141)
Internet			-0.019**
			(0.009)
Constant	4.618***	0.501	-0.932
	(0.021)	(0.802)	(1.675)
Year FE	No	Yes	Yes
Firm FE	No	Yes	Yes
Observations	23797	23797	23797
Adj_R^2	0.038	0.742	0.742

*Note: Standard errors reported in parentheses are clustered at the firm level, and standard errors are provided in parentheses. ***, **, and * denote the statistical significance of a two-tailed test at the 1 %, 5 %, and 10 % levels respectively.*

Robustness Tests

Parallel Trend Tests

The validity of the difference-in-differences (DID) approach hinges on the parallel trend assumption (Zhang & Sok, 2025). That is, a key prerequisite for identifying causal effects using DID models is that there should be no systematic differences in pre-treatment trends between the treatment and control groups. In the context of this study, the parallel trend assumption implies that, in the absence of the establishment of big data administration agencies, firms in treated and untreated cities would have exhibited similar trends in upgrading over time. To empirically assess whether this assumption holds, we conduct a parallel trend test based on Model (2), specified as follows:

$$Upgrade_{i,t} = \alpha + \sum_{t=-5}^{6} \delta_t D_{it} + \varphi Fc_{i,t} + \rho Cc_{c,t} + \mu_i + \lambda_t + \varepsilon_{i,t} \quad (2)$$

In this analysis, Dit represents a set of year-specific dummy variables equal to 1 if firm i is located in a city that established a big data administration agency in year t, and 0 otherwise. All other control variables are consistent with those specified in Equation (1). The primary focus is on the coefficient δ, which captures the differential in firm upgrading between treated and untreated cities across time. To further examine the validity of the parallel trends assumption, we conduct an event-study analysis using the year prior to policy implementation as the baseline period. Figure 1 presents the results of the parallel trends test based on the difference-in-differences framework. As shown, under the 95 % confidence interval, there are no statistically significant differences in firm upgrading levels between treated and untreated cities prior to the policy implementation. However, significant differences emerge after the establishment of big data administration agencies, providing strong evidence that the parallel trends assumption is satisfied and lending credibility to the baseline DID estimates.

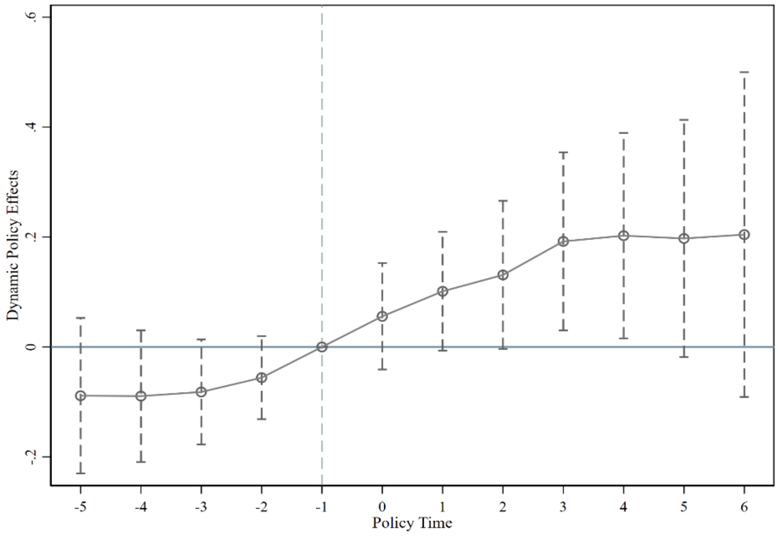

Figure 1. Parallel Trends Test for the Difference-in-Differences Model

Propensity Score Matching-Difference-in-Differences (PSM-DID) Estimation

To address potential endogeneity and sample selection bias, we conduct a robustness test using the PSM-DID approach. Although the establishment of big data administration agencies is largely exogenous, their spatial deployment may be influenced by local demands for data integration and sharing. In addition, firm heterogeneity and locational differences may introduce estimation bias. To mitigate these concerns, we use firm-level financial characteristics and regional attributes as matching covariates and estimate propensity scores via a logit model. Firms are then matched using year-by-year nearest-neighbor matching with caliper restrictions, and unmatched observations are discarded. The treatment and control groups are reconstructed accordingly. Column (1) of Table 5 reports the results from the PSM-DID estimation. After matching, the coefficient on Time×Treat remains significantly positive at the 5 % level, indicating that the establishment of big data administration agencies continues to have a significant and robust effect in promoting firm upgrading.

Controlling for Other Concurrent Policy Interventions

To ensure that the estimated effects are not confounded by overlapping policy interventions, we further control for three types of potentially concurrent policy shocks: (1) National Big Data Comprehensive Pilot Zones (Bigdata); (2) Information Benefiting People Pilot Cities (Information); (3) Smart City Pilot Programs (Smart). As shown in Columns (2) to (4) of Table 5, when each of the above policy dummies is included separately, the coefficient on Time×Treat remains significantly positive. In Column (5), when all three policy variables are included simultaneously, the core coefficient remains positively significant at the 10 % level. These results further confirm the robustness of the main finding: the establishment of big data administration agencies significantly enhances firm upgrading, independent of other parallel digital or information policy initiatives.

Table 5

Robustness Tests

	(1)	(2)	(3)	(4)	(5)
Variable	Upgrade	Upgrade	Upgrade	Upgrade	Upgrade
Time×Treat	0.098**	0.083**	0.076*	0.083**	0.077*
	(0.047)	(0.040)	(0.040)	(0.040)	(0.040)
Bigdata		0.331			0.349
		(0.358)			(0.359)
Information			0.098		0.095
			(0.061)		(0.061)
Smart				-0.057	-0.073
				(0.086)	(0.085)
Controls	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes
Observations	15542	23797	23797	23797	23797
Adj_R^2	0.787	0.742	0.742	0.742	0.742

*Note: The values in parentheses represent the robust standard errors for clustering at the firm level, and *, **, and *** indicate 10 %, 5 %, and 1 % significance levels, respectively.*

Placebo Test

To further rule out the influence of unobservable confounding factors, we conduct a placebo test by randomly assigning pseudo-treatment groups and fictitious policy implementation years. We repeat the simulation 1,000 times, re-estimating the DID regression each time. Theoretically, the placebo variable Time×Treat should exert no systematic effect, and the estimated coefficients should follow a normal distribution centered around zero. As shown in Figure 2, the simulated coefficients are tightly clustered around zero, while the baseline regression coefficient (0.082) is significantly distant from this distribution ($p<0.01$). These findings indicate that the observed effect of government data governance on firm upgrading is not driven by omitted variable bias or random shocks. The establishment of big data administration agencies exerts a statistically robust and significant impact on enhancing firm upgrading.

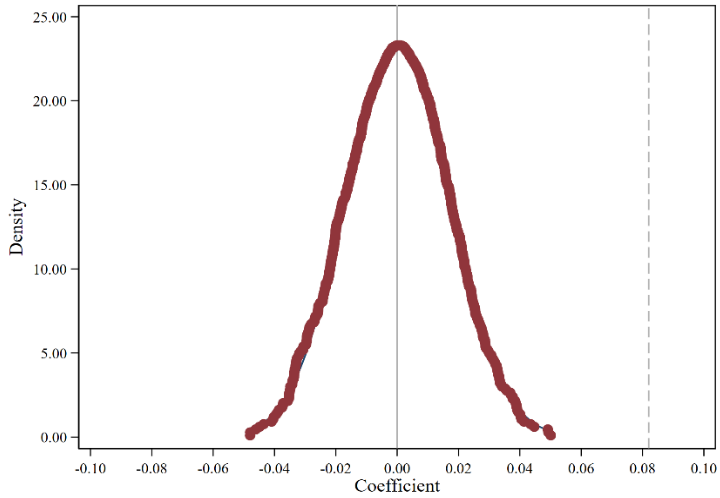

Figure 2. Placebo Test

Testing for Treatment Effect Heterogeneity in Multi-Period DID Models

Given that traditional two-way fixed effects (TWFE) models may suffer from treatment effect heterogeneity bias, we conduct further tests to assess the robustness of our baseline estimates under such conditions. First, we examine the share of negative weights in the estimation and find that 28.73 % of the weights are negative. This proportion is relatively low and suggests that the baseline DID estimates are reasonably robust. Second, we calculate the standard deviation of the estimator under heterogeneous treatment effects, which is 0.0436, a value close to zero. This indicates that some degree of instability may exist in the estimates.

To further validate the robustness of our main findings, we employ a series of heterogeneity-robust estimators, including the Stacked Regression Estimator, the imputation-based two-stage DID estimator, and the event-study adjusted DID estimator (Borusyak *et al.,* 2024; Gardner, 2022; Sun & Abraham, 2021). These alternative estimation strategies help correct for bias arising from staggered treatment adoption and non-parallel counterfactuals. Table 6 reports the estimation results. In all specifications, the coefficient on the policy interaction term remains positive and statistically significant, confirming that our previous findings are robust to treatment effect heterogeneity.

Table 6

Estimation Results Considering Heterogeneous Treatment Effects

Method	Estimated Treatment Effect Considering Heterogeneity
Imputation Estimator (did2s)	0.0002^{***} (0.000)
Group-Time Average Treatment Effect (eventstudyinteract)	0.1730^{*} (0.068)
Imputation Estimator (did_imputation)	0.2410^{**} (0.065)

*Note: The eventstudyinteract and did_imputation methods report policy effects that account for heterogeneous treatment effects at the group-time level. ***, **, and * denote the statistical significance of a two-tailed test at the 1 %, 5 %, and 10 % levels respectively.*

Addressing Endogeneity

(1) Instrumental Variable (IV) Estimation

A widely accepted approach to mitigating endogeneity bias is the use of a valid instrumental variable (IV) that satisfies both relevance and exogeneity conditions. Considering that government digital transformation heavily depends on the diffusion of digital infrastructure, we follow prior literature in constructing an infrastructure-based instrument.

Specifically, we use the interaction between the spherical distance from the firm's location to the nearest node city of the "Eight Vertical and Eight Horizontal" national fiber-optic backbone network and the lagged national number of broadband access ports as the instrument for government digital transformation. This is based on the rationale that digital government transformation relies on the deployment of digital technologies to promote policy transparency and government data sharing, thereby improving the public service environment for firms.

The closer a city is to a backbone network node, the more likely it is to benefit from high-speed internet connectivity and quality digital services, which in turn increases the likelihood that firms utilize digital government services. Thus, proximity to backbone network nodes positively correlates with a city's capacity and incentive to implement digital transformation, satisfying the relevance condition. Meanwhile, because the distance to the fiber-optic backbone network nodes is a historically determined geographic measure, it is plausibly exogenous and satisfies the exogeneity condition. As this geographic distance is time-invariant cross-sectional data, we follow the approach of prior literature (Nunn & Qian, 2014). To construct a time-varying instrument (Distance), we interact this variable with a time-varying national-level variable, specifically the lagged number of broadband ports nationwide.

Table 7 reports the results of the two-stage least squares (2SLS) estimation. Column (1) shows the first-stage regression, where the IV coefficient is significantly negative at the 1 % level, consistent with theoretical expectations. Column (2) presents the second-stage regression results. The Kleibergen-Paap LM F-statistic and Cragg-Donald Wald F-statistic are 216.236 and 370.087, respectively. Since both values are well above the Stock-Yogo critical value of 16.38 at the 10% significance level, this indicates no concern of weak instruments. The coefficient of Time×Treat in the second-stage regression remains positive and statistically significant, suggesting that even after addressing potential endogeneity concerns, the core findings of the baseline model remain robust.

(2) Addressing Omitted Variable Bias

To further mitigate potential bias arising from unobservable confounders, we implement several additional robustness checks to control for omitted variable bias.

First, we account for the influence of local industrial policy implementation. Given that different local governments may adopt heterogeneous industrial policies based on regional resource endowments and industrial structures, firms in different industries may face systematically different business environments, even when located within the same province. To address this issue, we include industry-province interaction fixed effects in the regression to account for such structural differences. Second, we control for macroeconomic fluctuations, which may vary across provinces due to differences in economic development stages and industrial composition. Economic cycles and changes in the global trade environment can affect provinces differently, leading to systematically different growth opportunities for firms located in different regions during the same year. To capture these dynamics, we add year-province interaction fixed effects to the baseline model. Third, to account for time-varying industry-level shocks, we introduce industry–year interaction fixed effects, which allow us to control for unobservable factors and potential heterogeneity across industries over time. As reported in Column (3) of Table 7, even after controlling for industry-year, province–industry, and year-province fixed effects, the coefficient on Time×Treat remains significantly positive. This confirms that our main findings are robust to a wide range of unobserved confounding factors.

Table 7

Addressing Endogeneity and Testing Policy Exogeneity

Variable	(1) First-stage Regression	(2) Second-stage Regression	(3) Controlling more Fixed Effects	(4) Exclude sub-provincial and provincial capital cities	(5) Assign a placebo policy implementation year
	Did	Upgrade	Upgrade	Upgrade	Upgrade
Distance	-0.005***				
	(-0.000)				
Time×Treat		2.706***	0.061*	0.095*	
		(0.658)	(0.035)	(0.051)	
Time2013×Treat					-0.008
					(0.216)
Controls	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes
Year×Industry	No	No	Yes	No	No
Province×Industry	No	No	Yes	No	No
Year×Province	No	No	Yes	No	No
Observations	23757	23757	23738	11372	6640
F-statistic	254.350				
LM F-statistic		216.236			
Wald F-statistic		370.087			
R^2/Center R^2	0.800	-0.201	0.827	0.702	0.807

*Note: The LM F-statistic refers to the Kleibergen-Paap LM F-statistic, and the Wald F-statistic refers to the Cragg-Donald Wald F-statistic. The values in parentheses represent the robust standard errors for clustering at the firm level, and *, **, and *** indicate 10 %, 5 %, and 1 % significance levels, respectively.*

Second, Testing for Policy Exogeneity. To ensure that our estimation results are not driven by policy selection bias, we implement two additional strategies to test for the exogeneity of the policy intervention. We exclude provincial capital cities and sub-provincial cities from the sample, retaining only cities with more comparable economic and social characteristics. The rationale is that larger administrative units may have systematically different capacities for digital governance. Column (4) of Table 7 reports the results of this restricted sample estimation. The coefficient on the key variable Time × Treat remains significantly positive, indicating that our baseline results are robust to sample selection based on administrative hierarchy. Then, we conduct a placebo policy timing test. Since the first batch of big data administration agencies was established in 2015, we restrict the sample to the pre-policy period (2011–2014) and assign 2013 as a fictitious policy implementation year, constructing a new interaction term Time2013×Treat. If the policy is truly exogenous, the coefficient on Time2013×Treat should be statistically insignificant. As reported in Column (5) of Table 7, the coefficient on the placebo interaction term is indeed not statistically significant, supporting the assumption of policy exogeneity and confirming that our core findings are not driven by endogenous policy assignment.

Third, Oster Test for Omitted Variable Bias. To further assess the potential influence of unobservable factors, we apply the (Oster, 2019) method for testing the robustness of coefficient estimates to omitted variable bias. Oster proposes two approaches to evaluate the impact of omitted variables on empirical results (Oster, 2019).

Method 1: Assume a value for δ, defined as the relative importance of unobservables compared to observables in explaining the outcome (typically set to $\delta=1$), and specify a value for the maximum R-squared (Rmax) of a hypothetical model that includes all omitted variables. Based on this, we simulate an adjusted coefficient estimate β_1^*. If β_1^* falls within the 95 % confidence interval of the baseline estimate β_1, the baseline result is considered robust. Method 2: Alternatively, given a value for Rmax and assuming $\beta_1 = 0$, we can calculate the implied δ. If $\delta > 1$, it indicates that omitted variables are unlikely to overturn the core empirical findings. In our empirical setting, we first augment the baseline regression (Equation 1) by including province–year fixed effects to control for all unobserved provincial-level shocks over time. This yields an adjusted R-squared greater than that of the baseline model (baseline R^2=0.7740), and we adopt this value as Rmax = 0.7812 for the Oster tests. The results are presented in Table 8. Row 1 reports the adjusted coefficient β_1^* under $\delta=1$. In both specifications, β_1^* falls within the 95 % confidence interval of the baseline estimate, confirming robustness. Row 2 presents the implied δ values under the assumption $\beta_1=0$. All δ estimates are greater than 1, suggesting that unobservable variables would need to be more influential than observed ones to eliminate the estimated effect, which is unlikely. Moreover, the small difference between Rmax (0.7812) and the baseline R^2 (0.7740) implies that the influence of unobserved province–year factors is minimal, further supporting the robustness of our baseline findings.

Table 8

Oster Robustness Test Accounting for Omitted Variables

Dependent Variable	R_{max}	Test Method	Evaluation Criterion	Estimated Result	Passes Test
Firm Upgrading	0.7812	Method 1	[0.0038， 0.1606]	β_1^*=0.0399	Yes
Firm Upgrading	0.7812	Method 2	δ>1	δ=1.8763	Yes

Further Analysis

Mechanism Tests

The preceding empirical results provide robust evidence that government data governance significantly promotes firm upgrading. But through what channels does this effect occur? As discussed in the theoretical analysis, government data governance may operate through two key mechanisms: optimizing the public service environment, and empowering firms' digital governance capabilities. In this section, we empirically investigate the validity of these proposed mechanisms.

Optimizing the Public Service Environment

As previously argued, government data governance plays a vital role in improving the quality of public services, which in turn facilitates firm upgrading. To more precisely identify this mechanism, we decompose the public service environment into three interrelated dimensions: government information transparency, digital infrastructure, and intellectual property (IP) protection. These dimensions capture distinct yet complementary aspects of the institutional conditions supporting enterprise upgrading.

In this analysis, each dimension of the public service environment is treated as a dependent variable, with government data governance as the key explanatory variable. All other controls and model specifications are consistent with the baseline regression in Equation (1). The estimation equation is specified as follows:

$$Environment_{c,t} = \theta_0 + \theta_1 Time \times Treat + \phi Cc_{c,t} + \lambda_c + \gamma_t + \sigma_{c,t} \quad (3)$$

First, government information transparency (Trans). Big data administration agencies typically disclose key public service information, such as approval procedures, required documentation, and processing timelines. We adopt the Fiscal Transparency Report of Chinese Municipal Governments published by Tsinghua University as a standardized and nationally recognized indicator for local transparency. This dataset reflects the extent to which local governments proactively release fiscal and administrative information, thereby reducing information asymmetry between the state and firms (Wan *et al.,* 2025).

Second, digital infrastructure conditions (Infrastructure). Drawing on the digital economy development index framework, we measure the quality of a city's digital development environment, a factor crucial for supporting firm upgrading (Luo *et al.,* 2023). We construct a composite digital infrastructure index using principal component analysis (PCA) based on the following variables: the number of internet users per 100 people, the proportion of employees in computer services and software, per capita telecom service volume, the number of mobile phone users per 100 people, and the Digital Inclusive Finance Index. These indicators are selected to jointly reflect network accessibility, digital workforce intensity, and digital service inclusiveness.

Third, intellectual property protection (Knowledge). The protection of innovation outcomes is a key element of a high-quality public service environment. We use the proportion of IP case closures in each city relative to the national total to proxy for the effectiveness of IP enforcement (Abdin *et al.,* 2024). This measure reflects local governments' administrative efficiency and legal capacity in safeguarding property rights. Higher IP closure rates indicate stronger legal protection and better innovation incentives.

Table 9 presents the DID estimation results based on city-level panel data. All models include city-level control variables as well as city and year fixed effects. The coefficients on the interaction term Time×Treat in Columns (1) to (3) are consistently positive and statistically significant, indicating that government data governance significantly promotes government information disclosure, improves digital infrastructure, and enhances IP protection, all of which contribute to creating a more enabling public service environment for firm upgrading.

Table

Mechanism I: Optimizing the Administrative Service Environment

	(1)	(2)	(3)
	Trans	Infrastructure	Knowledge
Time×Treat	3.920***	0.053**	0.088**
	(1.366)	(0.023)	(0.041)
Controls	Yes	Yes	Yes
City FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Observations	1860	2262	2153
Adj_R^2	0.599	0.934	0.793

*Note: The values in parentheses represent the robust standard errors for clustering at the firm level, and *, **, and *** indicate 10 %, 5 %, and 1 % significance levels, respectively.*

Empowering Firms' Digital Governance

The second mechanism examined in this study is how government data governance empowers firms' digital governance capabilities. Specifically, we assess this pathway through the lens of enterprise digital transformation:

First, we evaluate the use of data as a production factor by firms. We construct a proxy for this using Digital, the frequency count of digital transformation–related terms in firms' annual reports. This variable reflects the extent of digital engagement across dimensions such as artificial intelligence, big data, blockchain, cloud computing, and digital technology applications, following established text analysis approaches in the literature.

Second, we assess firm-level digital transformation using a Digital Transformation Index (Dindex). This index is compiled based on content extracted from listed companies' annual reports, fundraising disclosures, and certification announcements. It captures six dimensions: strategic leadership, technological drivers, organizational empowerment, digital application outcomes, and macro-level policy support. This multi-dimensional framework allows for a more comprehensive measurement of a firm's digital transformation.

Third, we analyze the proportion of digital assets using financial disclosure data. Following the literature, we calculate the ratio of digitally relevant intangible assets to total assets, using data disclosed in the notes to financial statements. This measure thereby captures the depth of digital asset accumulation.

Table 10 presents the DID estimation results for the impact of government data governance on the three measures above. All models include city-level and firm-level control variables, as well as firm and year fixed effects. The coefficients on the interaction term Time×Treat in Columns (1) to (3) are positive and statistically significant, indicating that government data governance significantly enhances enterprise digital transformation and empowers digital governance capabilities.

Table 10

Mechanism II: Empowering Corporate Digital Governance

	(1)	(2)	(3)
	Digital	Dindex	Dasset
Time×Treat	0.040*	0.294**	0.037*
	(0.022)	(0.147)	(0.021)
Controls	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes
Observations	23797	23797	19081
Adj_R^2	0.785	0.866	0.671

*Note: The values in parentheses represent the robust standard errors for clustering at the firm level, and *, **, and *** indicate 10 %, 5 %, and 1 % significance levels, respectively.*

Heterogeneity Analysis

To further explore the heterogeneous effects of government digital transformation on firm upgrading, we conduct subgroup analyses based on: (1) industry regulation intensity, (2) administrative level of the local big data administration agency, (3) whether the firm belongs to a high-tech industry, and (4) the level of firm data utilization.

Differences by Industry Regulation

Industries under strong government regulation tend to rely more heavily on state support and oversight, while non-regulated industries are generally more market-driven and innovation-led. Considering that government data governance may have asymmetric effects across industries with different regulatory intensities, we divide the sample into regulated and non-regulated industries and conduct subgroup regressions. Columns (1) and (2) of Table 11 report the heterogeneity results. The coefficient on Time×Treat is significantly positive at the 1 % level in the regulated industry subgroup, but becomes statistically insignificant in the non-regulated industry subgroup.

This suggests that the positive impact of government data governance on firm upgrading is more pronounced in regulated industries. One possible explanation is that firms in regulated sectors typically face more complex operations, stricter compliance requirements, and a higher dependency on government data and guidance. The establishment of big data administration agencies in such contexts may provide more structured and accessible public data resources, thereby supporting better-informed decision-making and enhancing firms' competitiveness in these industries.

Differences by Administrative Level of Big Data Administration Agencies

In regions where big data administration agencies are established at higher administrative levels, these institutions often possess stronger technical capabilities, greater financial support, and more advanced human capital. As a result, such regions are better equipped to provide enterprises with robust data infrastructure and technical services, thereby enhancing firms' capacity and motivation for upgrading. In this study, we classify big data agencies that are directly under provincial governments and hold department-level (Zhengtingji) administrative status as high-level institutions. Agencies with other administrative affiliations are categorized as lower-level institutions. We divide the full sample accordingly into high-level and low-level regions and conduct subgroup regressions. Columns (3) and (4) of Table 11 report the results. The coefficient on Time×Treat is positive and statistically significant at the 1 % level in high-level regions, while it is statistically insignificant in low-level regions.

These findings suggest that the positive effect of government digital transformation on firm upgrading is concentrated in regions with higher-level digital governance institutions. A plausible explanation is that governments in such regions tend to be more proactive and effective in the formulation and implementation of digital policies. They are also more responsive to firm needs and more capable of introducing and enforcing supportive measures that facilitate enterprise innovation and transformation, ultimately resulting in stronger firm upgrading outcomes.

Table 11

Heterogeneity Tests

	(1)	(2)	(3)	(4)
	Regulated Industries	Non-Regulated Industries	Higher-Level Institutions	Lower-Level Institutions
	Upgrade	Upgrade	Upgrade	Upgrade
Time×Treat	0.132^{**}	0.034	0.204^{***}	0.028
	(0.063)	(0.043)	(0.066)	(0.046)
Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
N	8990	14704	7381	16404
Adj_R^2	0.733	0.773	0.730	0.752
P-statistic	0.000		0.000	

*Note: The values in parentheses represent the robust standard errors for clustering at the firm level, and *, **, and *** indicate 10 %, 5 %, and 1 % significance levels, respectively.*

High-Tech vs. Non-High-Tech Industries

Firms in high-tech industries tend to possess stronger innovation capabilities and a higher capacity to adopt and integrate emerging technologies into their operations. Therefore, they are expected to be more responsive to government digital transformation initiatives. Based on industry classification, we divide the full sample into high-tech and non-high-tech industry subgroups and perform separate regressions. Columns (5) and (6) of Table 12 present the results. In the high-tech industry subgroup, the coefficient on Time × Treat is positive and statistically significant at the 1 % level. In contrast, the coefficient is not statistically significant in the non-high-tech industry group.

These findings indicate that the positive effect of government digital transformation on firm upgrading is primarily concentrated in high-tech industries. High-tech firms typically accumulate large volumes of data and maintain advanced digital infrastructure. Operations in R&D, product development, and market forecasting rely heavily on timely, high-quality data. This highlights a strategic reliance on data within these specific functions. Big data administration agencies can provide these firms with large-scale, high-quality public data, which enhances their innovation efficiency and adaptability to market conditions. By contrast, firms in non-high-tech industries are more market-driven and less dependent on government data, which may explain the weaker observed impact of digital transformation policies on their upgrading behavior.

Differences by Level of Data Factor Utilization

We construct a firm-level measure of data factor utilization based on the frequency of disclosures related to five categories of digital technologies in annual financial reports: artificial intelligence, blockchain, cloud computing, big data infrastructure, and big data applications. The total number of mentions across these categories is summed to reflect the intensity of data factor usage. Firms are then divided into two subgroups based on the annual median of this composite indicator: one with high data factor utilization and one with low data factor utilization. We conduct subgroup regressions accordingly. Columns (7) and (8) of Table 12 present the results. For firms with higher levels of data factor usage, the coefficient on Time×Treat is positive and significant at the 10% level. In contrast, for firms with lower levels of data factor usage, the coefficient is statistically insignificant.

These findings suggest that the effect of government digital transformation on firm upgrading is stronger for firms with more intensive data usage. Such firms often possess greater data processing capabilities and more advanced digital infrastructure due to their internal development needs. When government big data administration agencies provide richer public data resources, these firms are better positioned to rapidly integrate and leverage this information. This enhances their decision-making efficiency and accelerates upgrading. By comparison, firms with low data factor usage tend to have weaker data-driven innovation capabilities and lower digital awareness. Their slower adaptation to new technologies may hinder their ability to capitalize on the services and infrastructure provided by government digital governance institutions.

Table 12

Heterogeneity Tests

Variable	(1)	(2)	(3)	(4)
	High-Tech Industries	Non-High-Tech Industries	Firms with High Data Factor Utilization	Firms with Low Data Factor Utilization
	Upgrade	Upgrade	Upgrade	Upgrade
Time×Treat	0.174***	0.045	0.100*	-0.008
	(0.067)	(0.042)	(0.058)	(0.048)
Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Observations	7634	16059	14653	8709
Adj_R^2	0.782	0.731	0.760	0.787
P-statistic	0.000		0.040	

*Note: The values in parentheses represent the robust standard errors for clustering at the firm level, and *, **, and *** indicate 10 %, 5 %, and 1 % significance levels, respectively.*

To formally test the differences in coefficients across subgroups, we conduct Fisher's permutation test using 1,000 bootstrap resamples. The resulting p-values for the Time×Treat coefficient confirm the robustness and statistical significance of the subgroup regression results reported above.

Conclusion and Policy Implications

Understanding how government governance affects firm-level decision-making is not only a critical theoretical entry point for examining the relationship between the state and the market but also a topic of sustained academic interest. However, the role of government data governance, as an innovative mode of public sector reform, has received limited empirical attention in the existing literature, particularly regarding its impact on firm upgrading and the underlying mechanisms. This study exploits the policy implementation of municipal big data administration agencies as a quasi-natural experiment, and applies a difference-in-differences (DID) approach to systematically evaluate the effect of government data governance on firm upgrading. The results show that government data governance significantly promotes firm upgrading, and this finding remains robust across a series of sensitivity checks. Heterogeneity analysis reveals that the upgrading effect is more pronounced among: firms in heavily regulated industries, firms in high-tech sectors, firms with higher levels of data utilization, and regions where big data agencies are established at higher administrative levels. Mechanism analysis further demonstrates that government data governance facilitates firm upgrading through two key pathways: optimizing the public service environment, and empowering firms' digital governance capabilities.

Based on the empirical findings, the following policy implications are proposed:

First, governments should strengthen data governance practices that directly correspond to the two mechanisms identified in this study, namely optimizing the public service environment and empowering firms' digital governance capabilities. Based on the mechanism analysis showing that government data governance improves information transparency, digital infrastructure, and intellectual property protection, policymakers should prioritize the development of open data platforms, cross-departmental data-sharing mechanisms, and IP enforcement systems to foster a more innovation-friendly market environment. Moreover, governments should update governance models by integrating next-generation technologies such as big data, cloud computing, and artificial intelligence into public management to build a transparent and efficient digital governance framework. To enhance policy feasibility, these initiatives can be advanced through phased implementation and public–private collaboration, ensuring that investment costs are balanced with long-term efficiency gains.

Second, firms should align their upgrading strategies with government data governance initiatives and actively enhance their internal digital governance capabilities. In line with the empirical findings that firms benefit more from data governance in regions with stronger digital infrastructure and higher levels of data utilization, enterprises should invest in digital transformation, adopt intelligent data management systems, and strengthen data analytics capabilities to improve efficiency and innovation performance.

Third, local governments should implement differentiated and targeted policies based on firm characteristics and industry heterogeneity revealed in the study. The results show that the positive impact of government data governance is more pronounced among high-tech industries, regulated sectors, and firms with higher digital readiness. Accordingly, policymakers should design tailored incentives such as digital transformation subsidies, pilot projects for industrial data platforms, and customized IP protection services to accelerate regional digital transformation and promote sustainable industrial upgrading. Furthermore, establishing regular policy evaluation mechanisms can improve cost–benefit efficiency and ensure the sustained effectiveness of these differentiated measures.

Limitations and future research

This study demonstrates that government data governance significantly promotes firm upgrading through optimizing the public service environment and empowering firms' digital governance capabilities. However, several limitations should be acknowledged. First, the sample focuses on Chinese A-share listed firms, which may affect the representativeness of the results. Second, the analysis mainly captures short- and medium-term effects, while the long-term impact of data governance requires further examination. Third, issues such as data privacy, information security, and the digital divide were not explored in depth and deserve future attention. Finally, cross-country comparative studies could provide valuable insights into how different governance contexts influence the effectiveness of data governance in promoting firm upgrading.

Funding

This research was funded by the National Social Science Fund General Project of China, grant number 25BGJ057, the Innovation Project of Guangxi Graduate Education of China, grant number YCBZ2024021, the National Statistical Science Research Major Project of China, grant number 2024LD003.

References

Abdin, J., Sharma, A., Trivedi, R., & Wang, C. (2024). Financing constraints, intellectual property rights protection and incremental innovation: Evidence from transition economy firms. *Technological Forecasting and Social Change*, 198, 122982. https://doi.org/10.1016/j.techfore.2023.122982

Ali, M. A., Hoque, M. R., & Alam, K. (2018). An empirical investigation of the relationship between e-government development and the digital economy: The case of asian countries. *Journal of Knowledge Management*, 22(5), 1176–1200. https://doi.org/10.1108/JKM-10-2017-0477

Borusyak, K., Jaravel, X., & Spiess, J. (2024). Revisiting event-study designs: Robust and efficient estimation. *Review of Economic Studies*, 91(6), 3253–3285. https://doi.org/10.1093/restud/rdae007

Cheng, X., Chen, J., & Sheng, Y. (2025). Digital government construction and stock price synchronicity: Evidence from China. *Pacific-Basin Finance Journal*, 90, 102688. https://doi.org/10.1016/ j.pacfin.2025.102688

Gardner, J. (2022). Two-stage differences in differences (No. arXiv:2207.05943). arXiv. https://doi.org/10.48550/arXiv.2207.05943

Grimaldi, M., Greco, M., & Cricelli, L. (2021). A framework of intellectual property protection strategies and open innovation. *Journal of Business Research*, 123, 156–164. https://doi.org/10.1016/j.jbusres.2020.09.043

Li, X., & Yue, S. (2025). Does the government digital attention improve China's digital economy output efficiency: Accelerator or inhibitor. *Economic Analysis and Policy*, 85, 607–625. https://doi.org/10.1016/j. eap. 2024.12.020

Liang, Y., & Lu, W. (2024). Do export tax rebates promote firm upgrading? Evidences from labor-intensive firms in China. *Journal of Asian Economics*, 90, 101690. https://doi.org/10.1016/j.asieco. 2023.101690

Lo, K. L., Zhang, J., & Xia, F. (2023). Does digital technology innovation work better for industrial upgrading? An empirical analysis of listed Chinese manufacturing firms. Applied Economics Letters. https://www.tandfonline.com/doi/abs/10.1080/13504851.2022.2098238

Luo, S., Yimamu, N., Li, Y., Wu, H., Irfan, M., & Hao, Y. (2023). Digitalization and sustainable development: How could digital economy development improve green innovation in China? *Business Strategy and the Environment*, 32(4), 1847–1871. https://doi.org/10.1002/bse.3223

Ma, S., & Appolloni, A. (2025). Can financial flexibility enhance corporate green innovation performance? Evidence from an ESG approach in China. *Journal of Environmental Management*, 387, 125869. https://doi.org/10.1016/j.jenvman.2025.125869

Ma, S., Zeng, H., & Abedin, M. Z. (2025). The impact of the reforms in the Chinese equities exchange and quotations on innovation in cross-border e-commerce enterprises. *Asia Pacific Business Review*. https://www.tandfonline.com/doi/abs/10.1080/13602381.2025.2510509

Mergel, I. (2019). Digital service teams in government. *Government Information Quarterly*, 36(4), 101389. https://doi.org/10.1016/j.giq.2019.07.001

Nunn, N., & Qian, N. (2014). US food aid and civil conflict. *American Economic Review*, 104(6), 1630–1666. https://doi.org/10.1257/aer.104.6.1630

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. Journal of *Business & Economic Statistics*, 37(2), 187–204. https://doi.org/10.1080/07350015.2016.1227711

Park, S., & Gil-Garcia, J. R. (2022). Open data innovation: Visualizations and process redesign as a way to bridge the transparency-accountability gap. *Government Information Quarterly*, 39(1), 101456. https://doi.org/10.1016/j.giq.2020.101456

Ripamonti, J. P. (2024). Does being informed about government transparency boost trust? Exploring an overlooked mechanism. *Government Information Quarterly*, 41(3), 101960. https://doi.org/10.10 16/j.giq.2024.101960

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2), 175–199. https://doi.org/10.10 16/j.jeconom.2020.09.006

Verhoogen, E. (2023). Firm-level upgrading in developing countries. *Journal of Economic Literature*, 61(4), 1410–1464. https://doi.org/10.1257/jel.20221633

Wan, X., Guo, M., Jiang, P., & Xu, L. (2025). Government accounting reform, fiscal transparency, and local government debts: Evidence from China. *Pacific-Basin Finance Journal*, 91, 102763. https://doi.org/10.1016/j.pacfin.2025.102763

Wang, X., & Yang, G. (2025). Does being embedded in RTA networks promote firm productivity? Evidence from Chinese firms. *International Review of Economics & Finance*, 99, 103973. https://doi.org/10.1016/j.iref.2025.103973

Zeng, H., Liu, H., Yan, H., & Ma, S. (2025). Biodiversity risk and global stock markets: A cross-national heterogeneity analysis based on quantile-on-quantile methods. *Borsa Istanbul Review*, 25(6), 1518–1529. https://doi.org/10.1016/j.bir.2025.10.013

Zhang, L., & Sok, S. (2025). The impact of government open data platform construction on corporate capital market performance: Evidence from stock liquidity. *Pacific-Basin Finance Journal*, 90, 102667. https://doi.org/10.1016/j.pacfin.2025.102667

Zhou, Y., Zhu, S., & He, C. (2019). Learning from yourself or learning from neighbours: Knowledge spillovers, institutional context and firm upgrading. *Regional Studies*. https://www.tandfonline.com /doi/abs/10.1080/00343404.2019.1566705

Zhu, M., & Tomasi, C. (2020). Firms' imports and quality upgrading: Evidence from Chinese firms. *The World Economy*, 43(5), 1371–1397. https://doi.org/10.1111/twec.12934

Zhu, Y., & Yu, D. (2024). Can government digitalization promote firm productivity? Evidence from Chinese listed firms. *Applied Economics,* 56(33), 3939–3952. https://doi.org/10.1080/00036846. 2023.2208853

Authors' Biographies

Dan Wang is a Ph.D. candidate in the School of Economics at Guangxi University, China. Her research interests are in the areas of corporate finance and related fields. Her scholarly works have been published in prestigious journals including Technology in Society, Theoretical and Applied Climatology, Environmental Engineering & Management Journal (EEMJ). ORCID: https://orcid.org/0009-0007-8172-5669

Jian Hu is an Assistant Researcher at the School of Statistics, Tianjin University of Finance and Economics, China. His primary research domains include institutional economics and corporate finance. He has published extensively in internationally recognized journals such as the International Review of Financial Analysis, Technology in Society, Theoretical and Applied Climatology, Journal of Cleaner Production, Environmental Engineering & Management Journal (EEMJ), Global NEST Journal, Energy Exploration & Exploitation. ORCID: https://orcid.org/0009-0003-6691-2360

Shenglin Ma is a Candidate of North University of China whose research focuses on cross-border e-commerce. His paper have been published in journals such as Applied Economics, International Review of Economics & Finance, International Review of Financial Analysis, Finance Research Letters, Sustainable Cities and Society and Asia Pacific Business Review. ORCID: https://orcid.org/0009-0008-9517-2775

The article has been reviewed.
Received in June 2025; accepted in February 2026.