

From data factors to new quality productivity: Spatial and governance pathways of China's industrial transformation

Meihua Qiao

Liaocheng University

Jiaqi Xu

15063292610@163.com

Liaocheng University

Zhe Li

Liaocheng University

Article

Keywords: Data elements, new quality productivity, government attention, spatial spillovers

Posted Date: February 3rd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-8100262/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

New quality productivity represents an advanced form of industrial productivity driven by advanced technology, efficiency, and quality. In China, industrial upgrading remains constrained by reliance on traditional production paths and an uneven digital transformation. This study examines how data elements, emerging as a new factor of production, affect the development of new quality productivity across 30 Chinese provinces from 2013 to 2023. Using two-way fixed effects, panel thresholds, and the spatial Durbin model, the analysis shows that data elements significantly improve industrial productivity, with effects strengthening once government attention exceeds a critical threshold. Strong spatial spillovers reflect that data-driven development in one province stimulates productivity improvements in neighbouring regions. These findings indicate the intertwined roles of governance, spatial, and technology connectivity in shaping industrial transformation. This study offers new evidence for promoting inclusive, high-quality, and sustainable growth through digital capability building and policy coordination in the context of China's industrial modernization.

1. Introduction

New quality productivity is driven by scientific and technological innovation, which integrates new technologies, novel approaches, and high-quality talent. This paradigm not only disrupts traditional high-consumption and low-efficiency production models but also serves as a crucial force in promoting industrial transformation, upgrading industrial processes, and ensuring the high-quality development of the economy by improving technological capabilities, expanding innovation capacity, and optimising efficient resource allocation (Dai and Zheng, 2025). Fostering the new quality productivity of industrial enterprises is essential for reshaping a nation's industrial advantages and ensuring a robust national manufacturing sector. Industry reflects national strength and serves as the primary engine of economic growth. Currently, the Chinese industrial sector stands at a critical juncture, simultaneously facing an uphill climb and encountering hurdles to its growth. Despite achieving substantial developmental leaps over the past few decades, the Chinese industry continues to confront various structural contradictions and transformation challenges, with issues of economic imbalance and insufficient development remaining prominent (Xiong et al., 2023).

Data have emerged as a fundamental production factor alongside land, labour, capital, and technology, and play a strategic role in economic and social development (Wu and Li, 2024). The China ICT Academy's 2024 report projects that the scale of the Chinese digital economy will reach 53.9 trillion yuan in 2023, representing 42.8% of GDP. Data elements are characterised by incremental rewards and low-cost reuse (Mandler, 2024), enabling optimised resource allocation, empowering the real economy, and creating new opportunities for enhancing the quality of productivity in industrial enterprises. By the end of 2024, China had implemented more than 60 relevant policies, including the "Data Factor ×" Three-Year Action Plan, while 29 provinces have established special working groups to maximise the amplification, stacking, and multiplication effects of data factors. Nevertheless, the multiplier effect of data elements has yet to be fully realised: their contribution to industrial added value remains low, and their potential to enhance production efficiency and activate factor synergy has not been effectively translated into tangible outcomes. Consequently, it is imperative to elucidate the effects of data elements on the new-quality productivity of industrial enterprises and the mechanisms underlying these effects.

However, the existing research has three shortcomings. First, the research perspective is limited. Existing research mostly focuses on new quality productivity in agricultural and urban areas, and research on industrial new quality productivity is relatively insufficient. Academic studies have analysed the enabling effect of data elements. Second,

the existing research largely ignores the role of the threshold of government concern; the quantitative analysis of government concern is insufficient, and no attention has been given to the threshold of government involvement at which the enabling effect of data elements is optimised. Third, the spatial spillover mechanism through which data elements affect industrial new quality productivity is not yet clear.

The contributions of this paper can be divided into three main areas. First, takes on a new research perspective. In this paper, the new quality productivity of industrial enterprises is taken as the research perspective, an index system is constructed to measure the level of data elements and industrial new quality productivity in 30 provinces in mainland China from 2013 to 2023, and a two-way fixed-effects model is adopted to quantitatively analyse the direct impact of data elements on the new quality productivity of industrial enterprises. Second, the research content is innovative. In this paper, the impact of government involvement is quantitatively analysed. This paper uses the text word frequency method to extract data element-related words from government reports to quantify government involvement, builds a panel threshold model with this variable as the threshold variable, identifies the threshold level at which a difference in empowerment can be observed, and clarifies the mechanism underlying this threshold. Third, the research methodology is enriched. In this paper, the spatial Durbin model, which is based on provincial geography and economic linkages, is adopted to construct two types of matrices to analyse the differences in the spatial spillover effects of the data elements affecting industrial new quality productivity.

The remainder of the article is structured as follows. Section 2 discusses academic research addressing the existing issues and hypotheses. Section 3 outlines the relevant models and variables employed in this study. Section 4 presents and discusses the empirical findings. Finally, Section 5 summarises the principal conclusions of this research.

2. Literature review and hypothesis development

Previous studies have investigated data elements and new quality productivity, with the former focusing on the connotation of data elements, their measurement, and their impact on productivity, whereas the latter have focused on the connotations and characteristics of new quality productivity without considering data elements. The relationship between data elements and new quality productivity is mostly analysed theoretically; there is insufficient empirical evidence in the context of industrial enterprises, and there is disagreement on the impact of data elements on the new quality productivity of industrial enterprises. Therefore, this study examines the effect of data elements on the new quality productivity of industrial enterprises specifically.

2.1 Data elements

At present, the academic community has formed a fairly unified understanding of the connotation of data elements, which are generally regarded as an important material basis for the digital economy and industrial digital transformation, covering data collection, processing, application, and other aspects (Yang et al., 2023). In terms of the measurement of data elements, there are three main methods. One method is to define and quantify the level of data factor inputs through text analysis and other methods. The second method is to incorporate data factors into the production function and adopt econometric methods, such as the stochastic frontier production function, to dynamically assess their impact on total factor productivity. In the third method, data quality, system quality, and other indicators are combined to comprehensively evaluate the effectiveness and value of data factors (Fu et al., 2024; Hafner and da Silva, 2023; Motohashi et al., 2024).

With respect to the impact of data elements, this study focuses mainly on the path through which total factor productivity is enhanced in the manufacturing industry and reveals that data processing and application links play significant roles in enhancing technological efficiency, whereas the effect of pure data collection is limited. In addition, the impacts of data elements on corporate organisational performance, business innovation, digital economy development, fintech coupling, etc., have been assessed, with particular emphasis on the facilitating effect of data quality and data management on corporate decision-making and market competitiveness (Prashanth et al., 2024; Yang et al., 2025). On the whole, the impact of data elements is reflected mainly in the aspects of improving technical efficiency, optimising enterprise management, promoting industrial structure upgrades, and enhancing the vitality of the digital economy.

2.2 New qualitative productivity

Current research on new quality productivity focuses mainly on the connotations, characterisation, development status, and effects of new quality productivity. Scholars generally consider that new productivity emphasises technological innovation and sustainable development as the core; highlights the characteristics of digitalisation, intelligence, and sustainability; and is a new source of "kinetic energy" to promote high-quality development (Yue et al., 2024). New quality productivity is characterised by the innovation-driven, efficient allocation of factors, the upgrading of industrial structure, and green and low-carbon enterprises.

Studies on the effects of new productivity commonly describe its roles in promoting technological innovation, enhancing total factor productivity, promoting green development, and optimising industrial structure. The construction of digital infrastructure significantly promotes the development of new quality productivity, achieves a synergistic enhancement of the regional economy through promoting technological innovation, enhances the efficiency of factor allocation and industrial upgrading, and has spatial spillover effects (Dong and Liu, 2024). In specific industries, such as the sports industry, new quality productivity significantly promotes high-quality change through industry–university–research cooperation, and there is heterogeneity among different regions (Chen and Ye, 2025). In addition, new quality productivity plays a significant role in promoting regional green technological innovation, in which intellectual property protection plays an intermediary and amplifying role, but its promotional effect marginally decreases after reaching a certain level (Cheng et al., 2025). On the whole, new quality productivity not only enhances the innovation ability and production efficiency of enterprises and regions but also promotes green, low-carbon transformation and high-quality development, which is an important driving force for current economic and social transformation and upgrading.

2.3 Relationship between data elements and new quality productivity

The theoretical research on data elements and new quality productivity is abundant, but the existing empirical evidence focuses mostly on the fields of agriculture and listed companies, and the scientific measurement and evidence of the driving mechanisms at the level of industrial enterprises need to be deepened. Research on the relationship between data elements and new productivity is still based on theoretical analysis and logical reasoning and lacks a detailed measurement and identification of the mechanism underlying new productivity in industrial enterprises (Liu et al., 2025). Some of the studies involving the manufacturing industry also focus on single dimensions such as green innovation and total factor productivity; a systematic measurement framework and mechanism model have not yet been formed (Gao et al., 2023; Luo et al., 2025).

This study focuses on the industrial field, reviews the existing research and analyses the mechanism through which data elements affect the new quality productivity of industrial enterprises, and proposes research hypotheses. Then,

based on the data of 30 provinces (autonomous regions and municipalities directly under the central government) in mainland China from 2013–2023, an evaluation index system is developed, and the development level of the new quality productivity of industrial enterprises and data elements is measured. Finally, a bidirectional fixed-effects model is used to assess the influence of data elements on the new quality productivity of industrial enterprises.

The effect of data elements on the new quality productivity of industrial enterprises remains unclear. Some scholars suggest that data factor inputs significantly increase total factor productivity and green productivity in industry (especially manufacturing), which is an important driver of new quality productivity in industry (Wang and Shi, 2024; Zhang et al., 2023). Related literature also explores the facilitating role of data elements through mechanisms such as innovation, resource allocation optimisation, digital transformation, etc., and empirical studies have also reported that data factors enhance the total productivity and green productivity of industrial firms by improving their new quality productivity. Empirical studies have also revealed that data elements promote firms' new quality productivity through enhancing the quality of information disclosure (Huang and Chen, 2024; Li and Li, 2025), improving innovation capability (Li et al., 2024), and optimizing supply chain efficiency (Xu et al., 2025; Zhong et al., 2025). However, other scholars point to structural constraints and data paradoxes, arguing that digitisation may limit technological efficiency and that some segments have limited effects on productivity enhancement or even negative effects due to resource mismatch and management complexity (Liu et al., 2025). Studies mention the risk of investment inhibition and industrial hollowing out; i.e., although digitization enhances green productivity, it inhibits fixed asset investment in the manufacturing industry and leads to industry chain out-migration (Wen et al., 2024), and the positive effects of data complexity and security issues may be weakened by mismanagement of data complexity and security issues (Zong and Guan, 2024).

On this basis, the panel threshold and spatial Durbin model are further adopted to thoroughly analyse the effect of data elements on the new quality productivity of industrial enterprises, providing empirical evidence for local policies.

2.4 Development of hypothesis

2.4.1 Direct effects analysis

The dialectical unity of social reproduction provides a theoretical foundation for the empowerment of data elements. Data elements fundamentally reshape the logic of economic operations by becoming deeply embedded in production, circulation, consumption, and distribution (Allioui and Mourdi, 2023). Among these links, the production link serves as the core domain for the synergy of three types of new quality labour elements. Industrial enterprises utilise production data to establish precise parameters for new quality labour tools, cultivate new quality labour through the development of digital skills, enhance new quality labour materials through data-driven approaches, and transition production from an experience-driven model to a data-driven model. This shift represents an extension of the application of new quality labour materials and offers feedback for optimising production factors. The data factor transcends traditional constraints on resources due to its noncompetitiveness and zero marginal cost (Acciarini et al., 2023). The data factor fosters the development of new quality labour materials and objects, necessitates the intelligent iteration of new quality labour tools, and facilitates the transformation of the labour force into a new quality labour force through the acquisition of digital skills (Yao et al., 2023). The dynamic adaptation of these three types of factors constitutes the core driving force for the manufacturing industry's transition to a data-driven paradigm. Data regulates the increase in entropy within the production system, enhances synergy, and ultimately establishes a self-organised system characterised by incremental rewards. Accordingly, the following research hypothesis was formulated:

H1: There is a positive correlation between the level of data factor development and the new quality productivity of industrial firms.

2.4.2 Analysis of nonlinear effects

The application of new technologies and new elements must be supported by appropriate infrastructure and innovation ecology (Wolfert et al., 2023). If the government's digital infrastructure investment is insufficient, 5G networks, industrial internet platforms, and other carriers are lagging, and the ability to collect, store, and analyse data elements is limited, the efficiency of the integration of data elements with the industrial production system will be directly restricted.

When the government's involvement reaches a specific level, it activates the empowering effect of data elements through a three-way mechanism. First, in terms of institutional supply, the government introduces policies such as a market-based allocation plan for data elements and a digital economy plan; optimises the institutional framework for data authentication, trading, and security; and reduces the cost of element circulation. Second, in terms of resource investment, the government has increased its engagement with new infrastructure and has constructed an area-wide 5G network and a regional industrial internet platform to provide material support for the efficient utilisation of data elements. Third, in terms of cultivating an innovation ecosystem, the government has established digital economy industrial parks and industrial support funds to promote the deep integration of data elements with AI and blockchain, forming a multiplier effect of innovation diffusion (Li and Wang, 2023), which confirms the logic of Schumpeter's innovation factor agglomeration to promote changes in the production mode. Based on the above analysis, the following research hypothesis was formulated:

H2: There is a nonlinear threshold characterising the promotional effect of the level of data factor development on the new quality productivity of industrial firms based on government involvement.

2.4.3 Spatial spillover effects

As the core production factor of the digital economy, the data factor, by virtue of noncompetition, high mobility, and strong permeability, has an impact on the new quality productivity of industrial enterprises that breaks through the spatial limitations of a single region, forming a significant spatial spillover effect. Data elements break geographical barriers through the internet and data platforms to realise the rapid dissemination of technology and knowledge across regions (Zhong and Zhang, 2024). For example, after the success of the data-driven intelligent production model of a regional enterprise, technological experience can diffuse to the periphery through the flow of talent and technological cooperation, and peripheral enterprises can learn to imitate and integrate into their own production, which is in line with the "E-curve" of innovation diffusion.

The flow of data elements can optimise regional industrial collaboration and supply chain linkages. After the core enterprises realise accurate prediction and intelligent scheduling of the supply chain through data, the upstream and downstream enterprises optimise production, inventory, and logistics through data sharing, which promotes intelligent transformation of the industrial chain as a whole and achieves systematic enhancement of new quality productivity.

The efficient use of data elements relies on digital infrastructures such as 5G networks and cloud computing centres, and the spatial radiation characteristics of such infrastructure hubs not only serve the local area but also improve the data circulation environment in the surrounding area, reduce the cost of data elements used by neighbouring enterprises, and indirectly improve their ability to collect, analyse, and apply data, which in turn

promotes the development of new quality productive forces (Fang et al., 2024). In summary, the following hypothesis is formulated:

H3: An increase in the level of data factor development in a region not only promotes an increase in the new quality productivity of local industrial enterprises but also positively affects the new quality productivity of industrial enterprises in neighbouring regions.

3. Methodology

3.1 Modelling

3.1.1 Benchmark regression model

Before the regression analysis, it was first necessary to perform a multicollinearity test on the variables, which resulted in a VIF value of 4.12, which is much lower than 10, indicating that there is no multicollinearity problem. Next, the Hausman test determined that a fixed-effects model was most appropriate.

Considering that the intensive introduction of national digital economy policies after 2020 may simultaneously affect the data factor development and new quality productivity of industrial enterprises in each province, the time fixed effects can eliminate the interference of such macro time trends and make the estimation results of the core variables less noisy. Therefore, a double fixed-effects model is adopted to analyse the impact of the level of data factor development on the new quality productivity of industrial enterprises according to the following expression:

$$Mqp_{it} = \beta_0 + \beta_1 Data_{it} + \beta_c Cv_{it} + \mu_i + v_t + \varepsilon_{it} \quad (1)$$

Where Mqp is the development level of the new quality productivity of industrial enterprises in province i in year t , and $Data$ is the development level of data factors in province i in year t . Cv_{it} denotes the control variable, μ_i denotes the province fixed effect, v_t stands for the time fixed effect, ε_{it} is the random error term, β_1 is the total effect of the data factors on the new quality productivity of industrial enterprises, and β_0 is the constant term. β_c represents the coefficients of the control variables on Mqp .

3.1.2 Panel threshold model

The panel threshold model is essentially a data-driven segmentation analysis tool, and its core role is to accurately capture the critical nodes and segmentation characteristics of the variable relationship, which not only solves the problem of the traditional model's inability to explain the nonlinear relationship but also provides quantitative support for understanding the mechanism of the variable's role and formulating precise policies.

This model is selected in this study because it fits the characteristics of the possible marginal effect of data elements on the new quality productivity of industrial enterprises, can quantify the role of the differences between the different data element development level zones, can provide a theoretical basis for differentiated policy formulation, and can comprehensively reveal the nonlinear mechanism of the data elements driving the new quality productivity of industrial enterprises.

Therefore, combining the theoretical analysis, the model sets a single heavy threshold, dividing the sample into two intervals, and the core formula is as follows.

$$Mqp_{it} = \beta_{01} + \beta_{11}Cv_{it} + \beta_{x1}Data_{it} + \mu_i + \varepsilon_{it}, Dep_{it} \leq \gamma_1 \quad (2)$$

$$Mqp_{it} = \beta_{02} + \beta_{12}Cv_{it} + \beta_{x2}Data_{it} + \mu_i + \varepsilon_{it}, Dep_{it} > \gamma_1 \quad (3)$$

Where Dep_{it} is the threshold variable, which is the level of government involvement in province i in year t ; γ_1 is the threshold value; β_{01} and β_{02} are constant terms; β_{11} and β_{12} and β_{x1} and β_{x2} are the coefficients of control variables and $Data$, respectively; and the rest of the variables are defined as above. In practice, 500 bootstrap samples were used to calculate the standard errors and confidence intervals, and the significance of the thresholds was determined by the likelihood ratio test.

3.1.3 Spatial measurement model

The spatial Durbin model can not only quantify the spatial spillover effect of the explanatory variables on the response variables but also reveal the interaction effect of the levels of the explanatory variables in the spatially linked regions on the explanatory variables in the core region to characterise the spatial linkage between the variables more comprehensively. The formula is as follows:

$$Mqp_{it} = \rho \sum_{j=1}^n W_{ij} Mqp_{jt} + \varphi_1 Data_{it} + \varphi_2 Cv_{it} + \delta_1 \sum_{j=1}^n W_{ij} Data_{jt} + \delta_2 \sum_{j=1}^n W_{ij} Cv_{jt} + \mu_i + \varepsilon_{it} \quad (4)$$

Where W_{ij} denotes the spatial weight matrix, ρ denotes the spatial autoregressive coefficient, δ_1 denotes the level of development of the data element with the spatial interaction coefficient, δ_2 denotes the control variable with the spatial interaction coefficient, and the rest of the variable definitions are the same as those in the baseline regression model.

To select the spatial weighting matrix, a geographic matrix and an economic-geographic nested matrix that integrates geographic distance and GDP per capita (with geographic and economic dimensions weighted at 0.5 each) were selected for analysis, drawing on previous studies to ensure the reliability of the conclusions (Li and Liu, 2023; Li et al., 2024).

3.2 Description of the Variables

3.2.1 Explained variable

The system of indicators of the new quality productivity of industrial enterprises is shown in Table 1. The development level of new quality productivity (Mqp) of industrial enterprises is the explained variable of the model. Mqp consists of three main elements: "high-quality" workers, "new medium" labour materials, and "new material" labour objects (Yao et al., 2025). The new quality of productivity is the advanced productivity state in which innovation plays a leading role, which breaks from the traditional economic growth mode and productivity development path and is characterised by high technology, high efficiency, and high quality, in line with the new development concept (Xie et al., 2025). Therefore, this article refers to the existing research and constructs a comprehensive indicator system that includes process indicators of labourers, labour materials, and labour tools, as well as high-quality, high-tech, high-efficiency outcome indicators (Huang et al., 2023).

Table 1

Evaluation index system of the new quality productivity of industrial enterprises

First-level indicators	Second-level indicators	Third-level indicators	Indicator calculation
Labourer	High Technology	High-Tech Human Input	Full-time equivalent of R&D personnel in industrial enterprises above the designated size
	High Efficiency	Per Capita Income	Average wage of employed persons in urban manufacturing units
	High Quality	Employment Absorption	Employed persons in urban manufacturing units/Employed persons in urban units
Means of Labour	High Technology	Expenditure on New Product Development	Expenditure on new product development in industrial enterprises above a designated size
		R&D Expenditure	R&D expenditure in industrial enterprises above a designated size
		Patent Applications	Number of patent applications of industrial enterprises above the designated size
		Number of Valid Invention Patents	Number of valid invention patents of industrial enterprises above the designated size
	High Efficiency	Energy Consumption	Total industrial water consumption/Industrial added value
	High Quality	Pollution Control	Completed investment in industrial pollution control/Industrial added value
		Pollutant Emissions	Sulfur dioxide emissions/Regional gross domestic product
Object of Labour	High Technology	Number of New Projects	Number of new product projects of industrial enterprises above the designated size
		Finished Products	Finished products of industrial enterprises above the designated size
	High Efficiency	Sales Revenue of New Products	Sales revenue of new products of industrial enterprises above a designated size
		Asset-Liability Ratio	Total liabilities of industrial enterprises above the designated size/Total assets
		Current Asset Ratio	Total current assets of industrial enterprises above the designated size/Total assets
	High Quality	Scale Coordination	Number of large and medium-sized industrial enterprises/Number of industrial enterprises above the designated size
		Contribution Rate to Economic Growth	Industrial added value/Regional gross domestic product

3.2.2 Core explanatory variable

The development level of data elements (*Data*) is the core explanatory variable of the model. As shown in Table 2, the article refers to existing scholars' practices and selects 15 indicators from the three dimensions of data foundation support, data transformation ability, and data industry application, which are measured using the entropy method(Xie and Wang, 2023).

Table 2
Evaluation index system of the data elements

First-level indicators	Second-level indicators	Indicator calculation
Data infrastructure support	Optical fibre cable density	Length of Optical Fibre Cable Lines/Urban Area
	Internet broadband access density	Internet Broadband Access Ports/Year-end Permanent Population
	Mobile phone penetration rate	Mobile phone penetration rate
	Number of websites per 100 enterprises	Number of websites per 100 enterprises
	Number of domain names	Number of domain names
Data conversion capacity	Proportion of enterprises with E-commerce activities	Proportion of enterprises engaged in E-commerce transaction activities
	Total telecommunications business volume	Total telecommunications business volume
	Software business revenue	Software business revenue
	Express delivery business revenue	Express delivery business revenue
	Express delivery volume	Express delivery volume
Data industry application	E-commerce sales volume	E-commerce sales volume
	E-commerce procurement volume	E-commerce procurement volume
	Employed persons in urban units of information transmission, software, and information technology services	Employed persons in urban units of information transmission, software, and information technology services
	Number of Artificial Intelligence Enterprises	Number of Artificial Intelligence Enterprises
	Total index of digital inclusive finance	Total index of digital inclusive finance

3.2.3 Threshold variable

As the “visible hand”, the government plays a key role in activating the multiplier effect of data factors. Accordingly, in this article, government involvement is set as the threshold variable. The specific statistical methods are as follows: Python is used to process the government work reports of each province; first, we obtain the word frequency A of the keywords related to the digital economy and log them. To avoid measurement bias caused by word selection, the keywords are further screened, and the remeasured word frequency B is reregressed to ensure the robustness of the conclusions. Owing to space limitations, the statistical glossary is not shown and is kept for reference.

3.2.4 Control variables

Level of economic development (*Gdp*). Per capita national income is the causative factor of changes in industrial structure. Therefore, this article selects it as one of the control variables, which is measured by the per capita gross regional product published by the National Bureau of Statistics. The level of opening up to the outside world (*Open*). Expanding opening to the outside world can promote the expansion of export scale, which in turn drives the accumulation of material capital in the region, improves the efficiency of resource allocation, and ultimately promotes the accelerated upgrading of industrial structure. This study adopts the proportion of total imports and exports of the location of business units to regional GDP to measure the level of opening to the outside world. Urbanisation level (*Ur*). Urbanisation will drive industrial development in the direction of intensive recycling and innovation-driven transformation, which in turn will lead to the upgrading of traditional industries and the development of new, quality productivity among industrial enterprises. Population urbanisation is the core process of urbanisation, and the common indicator used by officials and researchers to measure the level of urbanisation is the proportion of the urban population to the resident population at the end of the year. Infrastructure (*II*). Infrastructure is an important material foundation and support condition for the development of new quality productivity among industrial enterprises, and its degree of perfection is directly related to the efficiency of factor flow and the level of industrial synergy. Therefore, this study takes the proportion of highway mileage in urban areas as the relevant variable. To alleviate the problem of heteroskedasticity, improve the effectiveness of the model estimation, and enhance the efficiency of the interpretation of economic impact, the data are logarithmized for all the variables.

3.3 Data sources and descriptive statistics

This study uses the sliding average of the values of the years before and after the indicator to replace the missing values of some indicators, and the descriptive statistics of the relevant variables are shown in Table 3.

Table 3
Descriptive statistics

Variable category	Variable	Symbol	Max	Min	Mean	Standard deviation	Obs.	VIF
Explained	New quality productive forces of industrial enterprises	Mqp	-0.131	-3.326	-2.241	0.656	330	-
Explanatory	Data element development level	Data	-0.210	-4.253	-2.464	0.773	330	4.63
Threshold	Government attention	Dep	4.317	0.00	2.863	0.677	330	1.58
Control	Economic development level	Gdp	12.208	10.003	10.980	0.440	330	9.01
	Opening-up level	Open	9.918	4.706	7.850	0.947	330	3.40
	Urbanization rate	Ur	-0.110	-0.971	-0.496	0.178	330	6.18
	Infrastructure level	Il	4.830	0.293	3.261	1.049	330	5.34

In this study, the panel data of 30 provinces (autonomous regions and municipalities directly under the central government) in mainland China (since there are many missing indicators for Tibet, the sample from Tibet is excluded) from 2013–2023 are used to conduct an empirical analysis, and the data are obtained mainly from the China Statistical Yearbook, China Industrial Statistical Yearbook, China Science and Technology Statistical Yearbook, and statistical yearbooks of each province (city).

4. Empirical analysis and discussion

4.1 Benchmark regression and robustness tests

The data elements contribute significantly to the new quality productivity of regional industrial firms. The results of the benchmark regression are shown in Table 4. Column (2) shows the results after the control variables are added, and the fitted coefficient of the data elements is 0.2261 and passes the significance test at the 5% level, which indicates that the conclusion is reliable and that $H1$ is verified.

The level of economic development and openness to the outside world are positively correlated with the new quality productivity of industrial enterprises, while the urbanisation rate and infrastructure development are negatively correlated. The reason behind this may be that regional economic strength provides material and institutional support for breakthroughs in new industrial productivity by guaranteeing investment in digital infrastructure, supplying R&D resources, and building a mature industrial ecosystem. Openness brings advanced manufacturing experience, and trade expansion releases market scale effects. In the deepening stage of urbanisation, the population is gathering in nonproductive service industries, and land resources are primarily utilised for urban construction, which leads to the contraction of labour and spatial factor supply in the industrial field. When traditional infrastructure essentially meets the needs of industrial infrastructure, excessive or homogenised infrastructure will crowd out innovation capital, leading to a decline in the efficiency of factor allocation and a shift from positive to negative marginal contribution.

Columns (3) through (5) report the results of robustness tests for lagged explanatory variables, lagged explanatory variables, and removal of municipalities, in that order. The results show that the signs of the coefficients of the lagged terms are consistent with those of the benchmark regression and that there is no substantial change in the level of significance, suggesting that the core findings are robust to the potential endogeneity problem.

Table 4
Benchmark regression and robustness test results

Variable	(1) Mqp	(2) Mqp	(3) Lagged Mqp	(4) Mqp	(5) Mqp
Data	0.2840** (0.0926)	0.2261** (0.0913)	0.3450*** (0.0981)	-	0.3498*** (3.3664)
Lagged data	-	-	-	0.1946** (0.0902)	-
Gdp	-	0.4495** (0.1855)	0.2397 (0.2036)	0.3806* (0.1990)	0.3946* (1.9665)
Open	-	0.2401*** (0.0347)	0.2213*** (0.0368)	0.2493*** (0.0354)	0.2067*** (5.7080)
Ur	-	-0.8429** (0.2784)	-0.7636** (0.3293)	-0.5407* (0.3259)	-2.0171*** (-4.9401)
Il		-0.2390*** (0.0520)	-0.2853*** (0.0588)	-0.2345*** (0.0571)	-0.3206*** (-5.7521)
Constant	-1.7231*** (0.1979)	-9.4710*** (2.2636)	-6.6789** (2.4717)	-8.7150*** (2.4172)	-7.1748*** (-3.0463)
Fixed effects	Controlled	Controlled	Controlled	Controlled	Controlled
Observations	330	330	300	300	286
Overall R-squared	0.9551	0.9653	0.9660	0.9710	0.5674

Note: *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Standard errors are in parentheses.

4.2 Endogeneity test

The systematic GMM estimation method is proposed due to the need to solve the endogeneity problem in dynamic panel models. The method performs better under finite sample conditions and can significantly reduce the estimation bias of the first-order difference GMM. Therefore, the system GMM estimation is adopted to address the endogeneity problem of the model, and the results are shown in Table 5.

The data elements remain significant at the 1% level, with a coefficient of 0.100 after overcoming the two-way causality and omitted variable interference, directly confirming their real drivers of new quality productivity in industrial firms. The reliability of the model is further strengthened by the test indicators: the p -value of AR(1) is 0.063, which suggests the existence of autocorrelation in the first-order difference residuals, implying that it is in line with the characteristics of the dynamic panel; the p -value of AR(2) is 0.177, which confirms that there is no second-order autocorrelation, which confirms the validity of the instrumental variables, and the Hansen test ($P = 0.664$) supports the joint validity of the instrumental variables. The regression coefficient of the lagged term of

industrial enterprises' new quality productivity is 0.982 and meets the significance level of 10%, which highlights the strong path dependence of industrial enterprises' new quality productivity, and the policy must coordinate intertemporal resources to maintain growth.

Table 5
Endogeneity test

Data	0.100** (0.0444)	Observations	300
		Adjusted R-squared	-
Lagged Mqp	0.982*** (0.0336)	Number of ID	30
		Year FE	YES
Gdp	0.0353 (0.0383)	AR(1)	z=-1.86 Pr > z = 0.063
Open	-0.0201 (0.0230)	AR(2)	z = 1.35 Pr > z = 0.177
Ur	0.0811 (0.100)	Sargan test	chi2(19) = 49.28 Prob > chi2 = 0.000
Il	0.0364* (0.0188)	Hansen test	chi2(19) = 15.90 Prob > chi2 = 0.664
Note: *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Standard errors are in parentheses.			

4.3 Threshold analysis

The results of the threshold regression are shown in Table 6. *Dep* has a coefficient of 0.0731. At a level of government involvement ≤ 2.3979 , the effect is not significant. When the level of government involvement is > 2.3979 , the level of development of data elements is significantly positive at the 10% level, the estimated *Dep* coefficient increases to 0.1062, and the role of data elements in promoting the new quality productivity of industrial firms is significantly strengthened. *Dep1* exhibits the same pattern, which further strengthens the robustness of the above findings, and *H2* is verified.

Table 6
Threshold regression results

Threshold variable	Threshold value	Threshold interval	Coefficient	t-value
Government Attention (Dep)	≤ 2.3979	First Threshold Interval	0.0731	1.31
	> 2.3979	Second Threshold Interval	0.1062*	1.83
Government Attention (Dep1)	≤ 1.9459	First Threshold Interval	0.0825	1.50
	> 1.9459	Second Threshold Interval	0.1241**	2.19
Note: *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.				

The reason for this is that when low policy attention is given, the role of policy as a catalyst for the release of the value of data elements is insufficient, and it is subject to the triple constraints of infrastructure, synergy, and technology. Enterprise investment is small, resulting in data that are difficult to scale for access, a lack of sharing mechanisms to form “data islands”, a lack of innovation power, and the inability to transform difficult-to-transform factors; therefore, the coefficient is small and not significant. In terms of policy attention above the threshold, resources are tilted to improve the infrastructure, the silos are broken, and a sharing environment is developed to promote the flow of data and to support R&D and personnel training to accelerate the transformation of factors. The available data on industrial enterprises helps to significantly promote the new quality productivity of industry and facilitates the synchronisation of relevant factors to accelerate these improvements.

Table 7 and Fig. 1 show the results of the threshold test and the regression results of the threshold effect using government involvement as the threshold variable, respectively. Government involvement measured in both ways shows a significant single threshold effect. In the threshold estimation test, the P value of *Dep* is 0.0040, the F value reaches 27.50, and the threshold estimate is 2.3979 after 500 bootstrap sampling simulations. The P value of *Dep1* is 0.0020, the F value is 32.14, and the threshold estimation is 1.9459, which strongly verifies the reliability of the threshold effect.

Table 7 Threshold estimation tests									
Threshold property	Threshold type	p-value	F-value	Bootstrap times	Critical value			Threshold estimate	95% confidence interval
					1%	5%	10%		
Dep	Single threshold	0.0040	27.50	500	22.43	14.41	11.24	2.3979	[2.1972, 2.4849]
Dep1	Single threshold	0.0020	32.14	500	24.23	15.40	12.26	1.9459	[1.7006, 2.0794]

4.4 Analysis of spatial spillover effects

4.4.1 Spatial correlation tests

In this article, the global Moran I index is used to test the spatial autocorrelation of data elements and industrial firms' new quality productivity in 30 provinces (autonomous regions and municipalities directly under the central government) in mainland China from 2013 to 2023, and the related variable selection matrix construction and data sources are described in detail in Part IV; thus, we do not repeat them here. The results are shown in Table 8.

Table 8
Global Moran's I index test results

Year	Data		Mqp	
	Geographical matrix	Economic matrix	Geographical matrix	Economic matrix
2013	0.1225*	0.1413**	0.2379***	0.1759***
2014	0.1470*	0.1491**	0.2164***	0.1513**
2015	0.1649**	0.1683***	0.2596***	0.1714***
2016	0.1673**	0.1884***	0.2356***	0.1541**
2017	0.1633**	0.1892***	0.2885***	0.1671***
2018	0.1687**	0.1668***	0.2833***	0.1654***
2019	0.1813**	0.1499**	0.2841***	0.1710***
2020	0.1902**	0.1315**	0.2928***	0.1725***
2021	0.2210***	0.1719***	0.2898***	0.1681***
2022	0.2290***	0.1748***	0.2854***	0.1679***
2023	0.2301***	0.1624***	0.2782***	0.1593**
<i>Note: *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.</i>				

There is spatial dependence between data factors and industrial firms' new quality productivity; i.e., areas with high levels of data factor development and industrial firms' new quality productivity are surrounded by clusters of areas with strong data factors and industrial firms with high levels of new quality productivity. Under both the geographic and economic matrices, the global Moran I index for the data factors is greater than zero and passes the test of significance at different levels in most years. The global Moran I index for the new quality productivity of industrial firms is also significantly positive. Specifically, under the geographic matrix, the Moran I index of the data elements in the period under investigation tends to increase overall, reflecting the gradual increase in the spatial agglomeration effect of the data elements in China, whereas the Moran I index of the new quality productivity of industrial enterprises is positive overall but fluctuates somewhat, reflecting a fluctuating change in its spatial agglomeration pattern.

4.4.2 Model identification and fitting results

The results of the above test reveal that both the data elements and the new quality productivity of industrial enterprises have spatial clustering characteristics; thus, it is necessary to introduce spatial factors into the analytical framework of the data elements affecting the new quality productivity of industrial enterprises. To select an appropriate spatial measurement model, **first**, this study carried out the LM test, as shown in Table 9. Both the spatial error model and the spatial autoregressive model can be used to analyse the spatial effect of data elements affecting the new quality productivity of industrial enterprises. **Second**, the Hausman test was applied to select the fixed effects. **Third**, the spatial Durbin model cannot be reduced to the spatial error model or the spatial autoregressive model with the help of the LR test.

Table 9
Selection of spatial measurement models

Test type	Test indicator	Geographical matrix	Economic matrix
LM test	LM-error	4.372 ^{***}	5.126 ^{***}
	Robust LM-error	3.123 [*]	22.483 ^{***}
	LM-lag	15.155 ^{***}	6.980 ^{***}
	Robust LM-lag	1.696 [*]	6.473 ^{**}
LR test	LR-spatial Error	Nondegenerate	Nondegenerate
	LR-spatial Error	Nondegenerate	Nondegenerate
<i>Note: *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.</i>			

The specific test results are shown in Table 10. From the empirical results, under both the geographic matrix and the economic-geographic nested matrix, the direction of the influence of the core variables on the new quality productivity of industrial enterprises and the significance characteristics are consistent: There is a significant positive effect of data elements on the new quality productivity of industrial enterprises. The coefficient under the geographic matrix is 0.1490 and is significant at the 10% level, while the coefficient under the economic matrix increases to 0.2252 and is significant at the 1% level, which is more prominent in terms of intensity and significance, indicating that data elements in regions with strong economic ties play a more significant role in directly empowering the new quality productivity of industrial enterprises. It is worth paying attention to the spatial spillover effect of the new quality productivity of industrial enterprises; the coefficient of this effect is 0.3274 under the geographic matrix, which is significant at the 1% level, and the coefficient is 0.3232 under the economic matrix, which also passes the significance test at the 1% level. These results clearly indicate that there is a significant positive spatial spillover characterising the effect of data elements on the new quality productivity of industrial firms, and the findings are robust.

Table 10
Results of the spatial econometric test

Variable	Geographical matrix	Economic matrix
Data	0.1490* (1.79)	0.2252*** (2.68)
Spatial lag of data	0.3658** (2.00)	0.1326 (0.54)
Spatial	0.3274*** (4.03)	0.3232*** (2.97)
sigma2_e	0.0137*** (12.73)	0.0143*** (12.75)
Control variables	YES	YES
Observations	330	330
R-squared	0.4612	0.4899
Note: *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Standard errors are in parentheses.		

4.4.3 Decomposition of Space Effects

The results of the decomposition of spatial effects based on the geographic and economic matrices show that there is distinct spatial heterogeneity in the impact of the level of development of data elements on the new quality productivity of industrial firms and that the characteristics of the role of the control variables vary according to the spatial correlation dimensions.

As shown in Table 11, under the geographic matrix, the direct effect of data elements is significantly positive, with a coefficient of 0.1784, passing the 10% significance test, which clearly demonstrates the direct role of the development of data elements in the region in promoting the new quality productivity of its own industrial enterprises. The indirect effect is particularly prominent, with a coefficient of 0.6057 and significance at the 5% level, which is also positive and significant, and the scale of the coefficient exceeds that of the direct effect, which means that the spatial spillover of geographic proximity is the core path through which data elements influence new productivity; i.e., the geographic coverage of data platforms and the physical accessibility of technological diffusion make it easier for neighbouring regions to share the dividends of the development of data elements. The total effect coefficient is 0.7841 and is significantly positive at the 1% level, which further verifies that data elements can promote the new productivity of industrial enterprises holistically through the combination of “local promotion + geographic overflow”.

In the economic matrix, the direct effect of data elements is significant, with a coefficient of 0.2381 and passing the 1% significance test, which exceeds the results of the geographic matrix both in terms of coefficient size and significance level. This finding illustrates that in regions with close economic ties (e.g., regions with similar GDP per capita), the data elements match more closely with the digitisation needs and application scenarios of local industries, and the direct driving effect is therefore stronger. However, the indirect effect coefficient of 0.3268 fails

the significance test, indicating that the spatial spillovers of data elements are more dependent on the “hard connection” of geographical proximity than the “soft connection” of economic isomorphism, which leads to a decrease in the significance of the total effect coefficient of 0.5648; this also leads to a decrease in the significance of the total effect coefficient of 0.5648. In summary, the core conclusions are consistent and stable under different spatial weight matrices with good robustness; i.e., Hypothesis H3 is verified.

Table 11
Spatial effects decomposition results

Variable	Geographical matrix			Economic matrix		
	Direct effect	Indirect effect	Total effect	Direct effect	Indirect effect	Total effect
Data	0.1784*	0.6057**	0.7841***	0.2381***	0.3268	0.5648
	(2.06)	(2.36)	(2.75)	(2.64)	(0.86)	(1.33)
Control variables	YES	YES	YES	YES	YES	YES
Observations	330	330	330	330	330	330
R-squared	0.4612	0.4612	0.4612	0.4899	0.4899	0.4899
<i>Note: *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Standard errors are in parentheses.</i>						

5. Conclusions

Giving full play to the advantages of data elements to increase the new quality productivity of industrial enterprises is highly important for the revitalisation of national industries. However, in reality, the role of data elements is not yet fully understood; theoretical studies are mostly from the perspective of agriculture and listed enterprises, and quantitative analysis is insufficient. Therefore, on the basis of a combination of relevant studies, this paper first elaborates on the mechanism and effect of data elements influencing the new quality productivity of industrial enterprises and proposes relevant research hypotheses. Second, the entropy value method is used to measure the development level of data and industrial enterprise data in China from 2013 to 2023. Finally, the baseline effect, threshold effect, and spatial spillover effect of the data factors on the new quality productivity of industrial enterprises are examined by using the two-way fixed-effects model, panel threshold model, and spatial Durbin model.

The main findings of this study are as follows:

First, there is a significant positive direct effect of data elements on the new quality productivity of industrial firms; i.e., the increase in the level of data elements helps to increase the level of new quality productivity of industrial firms, and the conclusions have passed three rounds of robustness tests of substitution variables, lagged variables, and elimination of extremes and have good reliability.

Second, there is a government involvement constraint on the positive impact of data elements on the new quality productivity of industrial firms. When government involvement does not cross the threshold, the impact of data elements on the new quality productivity of industrial firms is nonsignificant. Once this threshold is crossed, the

positive effect of data elements is significantly elevated. This reveals the key supportive role of the policy environment in unleashing the value of data elements.

Third, the impact of data elements on the new quality productivity of industrial enterprises is characterised by significant spatial spillovers. From the perspective of the geographic matrix, the indirect effect of data elements is greater than the direct effect, indicating that geographically neighbouring regions can share the dividends of data elements through technological diffusion and industrial chain synergy. From the perspective of the economic matrix, the direct effect of data elements is more significant, reflecting the strong spatial role of data elements in geographic proximity spillovers and the strong direct drive of economic linkages.

There are two points to be explored in future studies. First, owing to the difference in statistical caliber and missing data, some of the data were filled in with a sliding average, which affects the precision of the conclusion. Second, the tracking and integration of frontier theories of the digital economy are insufficient, and the adaptability is poor. Future research can continue to explore this topic in depth. First, the data sources can be expanded, or machine learning algorithms can be adopted to fill data gaps and reduce measurement bias. Second, the tracking of frontier theories of the digital economy should be strengthened, old and new theories should be integrated, and the complex reality should be accurately explained.

Declarations

Authors' contributions

MQ and JX composed the conception and design and drafted the article. MQ interpreted the data. JX and ZL collaborated on the study's writing. JX and ZL work on data methodology and analysis. MQ, JX, and XL made critical comments and amendments.

Funding statement

This study did not receive any funding.

Ethical approval

This study did not involve human participants; therefore, no ethical approval is required.

Informed consent

Not applicable.

Competing interest

The authors declare that they have no competing interests.

Data Availability Statement

Data will be available at the request of the authors.

References

1. Acciarini C, Cappa F, Boccardelli P et al (2023) *Rev Technovation* 123.
<https://doi.org/10.1016/J.TECHNOVATION.2023.102713>. How Can Organizations Leverage Big Data to Innovate Their Business Models? A Systematic Literature
2. Alloui H, Mourdi Y (2023) Exploring the Full Potentials of IoT for Better Financial Growth and Stability: A Comprehensive Survey. *Sensors* 23(19). <https://doi.org/10.3390/S23198015>
3. Chen XQ, Ye SZ (2025) Influence of New Qualitative Productivity Forces on Quality-Effective Change in China's Yangtze River Economic Belt Sports Industry: The Mediating Effect of Industry-University-Research Cooperation. *PLoS ONE* 20(2):e0319366. <https://doi.org/10.1371/JOURNAL.PONE.0319366>
4. Cheng K, Yin JH, Wang FY et al (2025) The Impact Pathway of New Quality Productive Forces on Regional Green Technology Innovation: A Spatial Mediation Effect Based on Intellectual Property Protection. *PLoS ONE* 20(4):e0319838. <https://doi.org/10.1371/JOURNAL.PONE.0319838>
5. Dai DB, Zheng Y (2025) The New Quality Productive Force, Science and Technology Innovation, and Optimisation of Industrial Structure. *Sustainability* 17(10):4439–4439. <https://doi.org/10.3390/SU17104439>
6. Dong CH, Liu DX (2024) How to Deeply Embed the Digital Economy into the Product Quality Upgrade? The Moderating Effect of Resource Allocation. *Kybernetes* 53(10):3875–3895. <https://doi.org/10.1108/K-11-2022-1619>
7. Fang YH, Wei W, Su RG (2024) Data Element Investment Strategy: How Can Leaders and Followers Innovate in Dynamic Market Environments. *PLoS ONE* 19(8):e0309659. <https://doi.org/10.1371/JOURNAL.PONE.0309659>
8. Fu Q, Nicholson GL, Easton JM (2024) Understanding Data Quality in a Data-Driven Industry Context: Insights from the Fundamentals. *J Ind Inf Integr* 42:100729–100729. <https://doi.org/10.1016/J.JII.2024.100729>
9. Gao Q, Cheng CM, Sun GL (2023) Big Data Application, Factor Allocation, and Green Innovation in Chinese Manufacturing Enterprises. *Technol Forecast Soc Change* 192.
<https://doi.org/10.1016/J.TECHFORE.2023.122567>
10. Hafner M, Mira da Silva M (2023) *Knowl Inf Syst* 66(3):1807–1831. <https://doi.org/10.1007/S10115-023-01994-4>. Towards a Taxonomy for Business Capabilities Determining Data Value
11. Huang YP, Chen ZG, Li HR et al (2023) The Impact of Digital Economy on Green Total Factor Productivity Considering the Labor-Technology-Pollution Factors. *Sci Rep* 13(1):22902–22902.
<https://doi.org/10.1038/S41598-023-50400-0>
12. Huang P, Chen XH (2024) The Impact of Data Factor-Driven Industry on the Green Total Factor Productivity: Evidence from China. *Sci Rep* 14(1):25377–25377. <https://doi.org/10.1038/S41598-024-77189-W>
13. Motohashi K, Koshiba H, Ikeuchi K (2024) Measuring Science and Innovation Linkage Using Text Mining of Research Papers and Patent Information. *Scientometrics* 129(4):2159–2179. <https://doi.org/10.1007/S11192-024-04949-W>
14. Li YX, Wang FS (2023) The Corporate Path to Green Innovation: Does the Digital Economy Matter? *Environ Sci Pollut Res Int* 30(32):79149–79160. <https://doi.org/10.1007/S11356-023-27818-0>
15. Li YY, Liu WJ (2023) Spatial Effects of Environmental Regulation on High-Quality Economic Development: From the Perspective of Industrial Upgrading. *Front Public Health* 11:1099887–1099887.
<https://doi.org/10.3389/FPUH.2023.1099887>
16. Li HJ, Zhang Y, Li Y (2024) The Impact of the Digital Economy on the Total Factor Productivity of Manufacturing Firms: Empirical Evidence from China. *Technol Forecast Soc Change* 207:123604–123604.
<https://doi.org/10.1016/J.TECHFORE.2024.123604>

17. Li WX, Cai JJ, Zhu YW et al (2024) Can Digital Finance Development Drive Green Transformation in Manufacturing? Evidence from China. *Environ Sci Pollut Res Int* 31(16):23876–23895. <https://doi.org/10.1007/S11356-024-32402-1>
18. Li J, Li ZJ (2025) Mechanisms of Corporate Digital Transformation on Asymmetric Capital Structure Adjustment—the Mediating Role of Information Asymmetry and Financial Stability. *Heliyon* 11(3):e41745–e41745. <https://doi.org/10.1016/J.HELİYON.2025.E41745>
19. Liu K, Chi MM, Zhou Q (2025) The Power Mechanism of China's Total Factor Productivity Improvement from the Perspective of New Quality Productivity: A Group Study Based on Dynamic fsQCA. *Sustainability* 17(6):2666–2666. <https://doi.org/10.3390/SU17062666>
20. Liu L, Xin YL, Liu B et al (2025) The Panel Threshold Analysis of Digitalization on Manufacturing Industry's Green Total Factor Productivity. *Sci Rep* 15(1):4336–4336. <https://doi.org/10.1038/S41598-025-86643-2>
21. Luo TY, Qu JJ, Cheng S (2025) Digital Transformation, Dynamic Capability and Total Factor Productivity of Manufacturing Enterprises. *Ind Manag Data Syst* 125(3):921–944. <https://doi.org/10.1108/IMDS-09-2024-0907>
22. Mandler M (2024) Increasing Returns and the Efficient Acquisition of Information. *J Econ Theory* 220:105815. <https://doi.org/10.1016/J.JET.2024.105815>
23. Prashanth M, Li HX, Helander N (2024) *Ind Manag Data Syst* 124(7):2364–2387. <https://doi.org/10.1108/IMDS-09-2023-0653>. Developing Data Analytics Capabilities: Integrating the Information Systems Success Model and the Resource-Based View
24. Wolfert S, Verdouw C, van Wassenae L et al (2023) Digital Innovation Ecosystems in Agri-Food: Design Principles and Organizational Framework. *Agric Syst* 204. <https://doi.org/10.1016/J.AGSY.2022.103558>
25. Wang XT, Shi XF (2024) Impact of Digital Transformation on Green Production: Evidence from China. *Heliyon* 10(15):e35526–e35526. <https://doi.org/10.1016/J.HELİYON.2024.E35526>
26. Wen HW, Liu YP, Liu YT (2024) Impact of Digitalization on Investment and Productivity of Manufacturing Industry: Evidence from China. *SAGE Open* 14(3). <https://doi.org/10.1177/21582440241281862>
27. Wu QY, Li SH (2024) Decarbonization by Digits: How Data Factors Drive Nonlinear Sustainable Dynamics in Manufacturing. *Appl Energy* 374:123967–123967. <https://doi.org/10.1016/J.APENERGY.2024.123967>
28. Xie WJ, Wang RX (2023) Application of Data Elements in the Coupling of Finance and Technology on the Digital Electronic Platform. *Electron Commer Res* 24(2):1435–1460. <https://doi.org/10.1007/S10660-023-09686-5>
29. Xie FS, Jiang N, Kuang XL (2025) Towards an Accurate Understanding of 'New Quality Productive Forces'. *Econ Polit Stud* 13(1):1–15. <https://doi.org/10.1080/20954816.2024.2386503>
30. Xiong X, Wang YX, Liu B et al (2023) The Impact of Green Finance on the Optimisation of Industrial Structure: Evidence from China. *PLoS ONE* 18(8):e0289844–e0289844. <https://doi.org/10.3390/SU17020633>
31. Xu YS, Wang RX, Zhang SF (2025) Digital Economy, Green Innovation Efficiency, and New Quality Productive Forces: Empirical Evidence from Chinese Provincial Panel Data. *Sustainability* 17(2):633–633. <https://doi.org/10.3390/SU17020633>
32. Yang YP, Chen N, Chen N et al (2023) The Digital Platform, Enterprise Digital Transformation, and Enterprise Performance of Cross-Border E-Commerce—From the Perspective of Digital Transformation and Data Elements. *J Theor Appl Electron Commer Res* 18(2):777–777. <https://doi.org/10.3390/JTAER18020040>
33. Yang P, Li H, Hou WJ et al (2025) The Analysis of Strategic Management Decisions and Corporate Competitiveness Based on Artificial Intelligence. *Sci Rep* 15(1):17942–17942. <https://doi.org/10.1038/S41598-025-02842-X>

34. Yao JF, Yang Y, Wang XC et al (2023) Systematic Review of Digital Twin Technology and Applications. *Visual Comput Ind Biomed Art* 6(1):10–10. <https://doi.org/10.1186/S42492-023-00137-4>
35. Yao L, Li AY, Yan EW (2025) Research on Digital Infrastructure Construction Empowering New Quality Productivity. *Sci Rep* 15(1):6645–6645. <https://doi.org/10.1038/S41598-025-90811-9>
36. Yue SS, Bajuri NH, Khatib SFA et al (2024) New Quality Productivity and Environmental Innovation: The Hostile Moderating Roles of Managerial Empowerment and Board Centralization. *J Environ Manag* 370:122423. <https://doi.org/10.1016/J.JENVMAN.2024.122423>
37. Zhang WJ, Xu N, Li CY et al (2023) Impact of Digital Input on Enterprise Green Productivity: Micro Evidence from the Chinese Manufacturing Industry. *J Clean Prod* 414. <https://doi.org/10.1016/J.JCLEPRO.2023.137272>
38. Zhong CB, Zhang C (2024) Can Data Elements Enhance Urban Innovation? Evidence from China. *China Econ Rev* 88:102307–102307. <https://doi.org/10.1016/J.CHIECO.2024.102307>
39. Zhong Y, Lai HS, Zhang L et al (2025) Does Public Data Openness Accelerate New Quality Productive Forces? Evidence from China. *Econ Anal Policy* 85:1409–1427. <https://doi.org/10.1016/J.EAP.2025.01.025>
40. Zong ZJ, Guan Y (2024) Transforming Knowledge, Innovation, and Efficiency. *J Knowl Econ* 16(1):1–40. <https://doi.org/10.1007/S13132-024-02001-Z>. AI-Driven Intelligent Data Analytics and Predictive Analysis in Industry 4.0

Figures

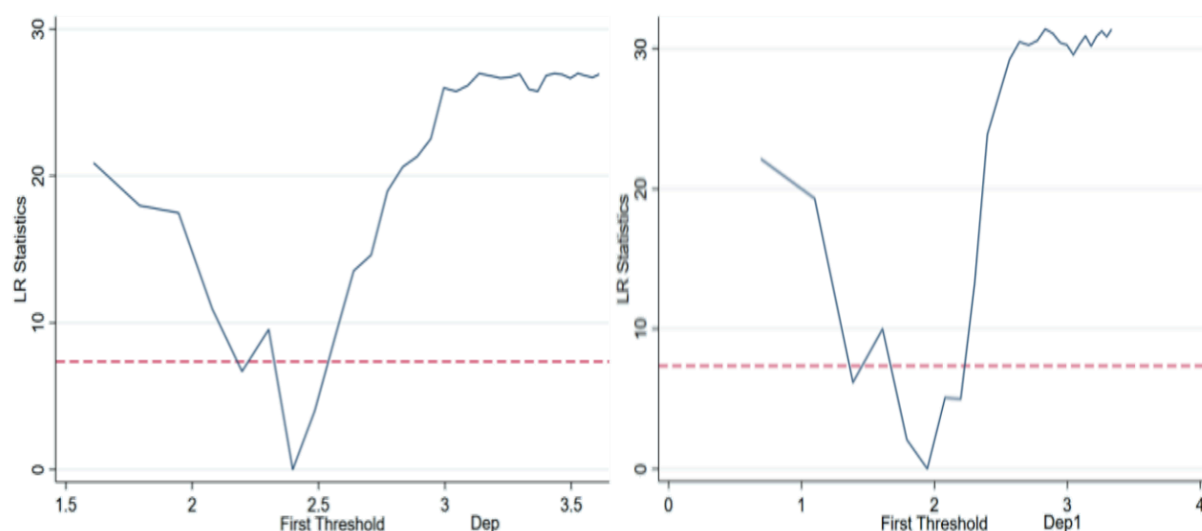


Figure 1

Threshold effect likelihood ratio statistic plot