Frontiers in Economics and Management Volume 7 Issue 11, 2026
ISSN: 2692-7608 DOI: 10.6981/FEM.202611_7(11).0006

The Effect of Data Factor Utilization Capability on Corporate Resilience

Yuxuan Li

School of Business, Nanjing Normal University, Nanjing 210023, China

Abstract

As global technological competition intensifies, firms' capability to use data as a factor of production may help strengthen their resilience. Using a sample of Chinese A-share listed firms from 2011 to 2024, this study investigates the effect of data factor utilization capability on corporate resilience and the mechanisms underlying this relationship. The results indicate that stronger data factor utilization capability significantly enhances corporate resilience, primarily through improved internal control and persistent innovation. Furthermore, the positive effect is more pronounced among firms operating in high-tech industries and those located in regions with more favorable innovation environments. These findings highlight the potential value of improving the institutional framework for data as a factor of production and supporting firms' effective use of data.

Keywords

Data Factors; Data Factor Utilization Capability; Corporate Resilience.

1. Introduction

In recent years, intensifying geopolitical conflicts and global economic uncertainty have substantially increased the external risks and internal management challenges confronting firms[1]. These shocks can disrupt business operations and spread through upstream and downstream relationships across industrial and supply chains, exposing and amplifying firm-level vulnerabilities. Strengthening corporate resilience has therefore become an increasingly important concern for corporate development.

In the digital economy, data have emerged as a new factor of production that contributes to value creation[2]. Data factor utilization capability concerns how effectively a firm uses data, rather than how much data it holds. It encompasses the acquisition, integration, and analysis of data for internal management, strategic decision-making, and other business activities[3,4]. When confronted with unexpected shocks, firms with stronger data factor utilization capability are better positioned to detect changes in market demand, production anomalies, and supply-chain risks and to adjust resource allocation and business strategies accordingly[1]. This capability thus offers a new avenue for strengthening corporate resilience.

Existing research has mainly examined how digitalization and digital technologies relate to corporate resilience. Digital technologies support data analysis, information processing, and informed decision-making during crises[5]. These capabilities can improve supply-chain coordination and business continuity, thereby supporting resilience[6-8]. As data have increasingly been recognized as a distinct factor of production, other studies have investigated the positive effects of open public data [9]and market-based allocation of data factors [10]on corporate resilience. However, this research focuses mainly on the external supply of data, rarely taking the process through which data generate value as its analytical starting point. Consequently, insufficient attention has been paid to whether firms can apply data effectively and translate that capability into greater resilience.

ISSN: 2692-7608 DOI: 10.6981/FEM.202611_7(11).0006

This study contributes to the literature in three ways. First, it examines corporate resilience through the capability to transform data resources into productive and operational capabilities. This perspective extends research on data as a factor of production and corporate resilience. Second, it examines internal control and persistent innovation as potential mechanisms linking data factor utilization capability to resilience. Third, it explores variation across industries and regional innovation environments to inform policies supporting firms' effective use of data.

2. Theoretical Analysis

2.1. Data Factor Utilization Capability and Corporate Resilience

Corporate resilience fundamentally refers to a firm's capacity to withstand adverse shocks and recover its development trajectory[5]. It depends on both available resources and the ability to recognize environmental changes and reorganize those resources. As both a carrier of information and a source of knowledge generation, data can support resilience by being embedded in the firm's "sensing–decision-making–reconfiguration" process[10]. Specifically, by analyzing data from multiple sources, including market demand, supply-chain operations, and technological developments, firms can promptly identify emerging risks and respond more quickly. In addition, sharing data across business functions also reduces departmental information barriers [2]and improves the timeliness and accuracy of data-driven decisions. When external shocks occur, firms can adjust operating strategies and identify alternatives to maintain operations and shorten recovery periods. These arguments motivate the following hypothesis.

H1: Stronger data factor utilization capability enhances corporate resilience.

2.2. Mechanisms Linking Data Factor Utilization Capability to Corporate Resilience

2.2.1. The Mediating Role of Internal Control.

Internal control is an important governance mechanism through which firms identify operational risks, standardize business processes, and safeguard strategy implementation. Traditional internal control relies heavily on fragmented information flows and retrospective inspections. Information delays, departmental silos, and managers' bounded rationality can limit its effectiveness. Greater application of data factors can improve the transparency and traceability of internal management information and encourage managers to emphasize ex ante risk alerts and ongoing monitoring, thereby improving the efficiency of risk identification and response[1]. High-quality internal control can also improve financial reporting quality[11], alleviate information concerns among external investors and financial institutions, and consequently reduce financing costs to some extent. This provides financial support for post-shock production recovery and resource restructuring. These arguments motivate the following hypothesis.

H2: Data factor utilization capability enhances corporate resilience by strengthening internal control.

2.2.2. The Mediating Role of Persistent Innovation.

Persistent innovation helps firms accumulate technological advantages and adapt to changing markets. By analyzing consumer demand, competitive dynamics, and internal R&D data, firms can identify promising innovation directions and R&D opportunities more accurately and allocate resources to projects with greater potential[12]. Data can also connect heterogeneous knowledge across internal departments as well as upstream and downstream partners, enabling firms to expand their knowledge networks and maintain continuity in R&D activities[10]. Technological accumulation and product pipelines developed through persistent innovation reduce dependence on a single technological pathway or source of market demand.

ISSN: 2692-7608 DOI: 10.6981/FEM.202611_7(11).0006

During shocks, firms with stronger data factor utilization capability may identify alternatives and restore growth more quickly[8]. Accordingly, the following hypothesis is proposed.

H3: Data factor utilization capability enhances corporate resilience by promoting persistent innovation.

3. Research Design

3.1. Model Specification

To examine and interpret the effect of data factor utilization capability on corporate resilience, the following baseline regression model is specified:

$$rir_{i,t} = \alpha_0 + \alpha_1 data_{i,t} + \gamma X_{i,t} + \delta_i + \lambda_t + \varepsilon_{i,t} \quad (1)$$

Here, i and t denote firms and years, respectively. The dependent variable, $rir_{i,t}$ represents the level of corporate resilience in firm i in year t ; $data_{i,t}$ is the core explanatory variable, measuring data factor utilization capability; $X_{i,t}$ is a vector of control variables; δ_i and $\lambda_t$denote individual (*firm*) and year fixed effects, respectively; and $\varepsilon_{i,t}$ is the error term. In the baseline regressions, standard errors are clustered at the firm level.

3.2. Variable Construction

Dependent variable. Corporate resilience captures a firm's capability to withstand adverse shocks and recover[5,13]. This study estimates total factor productivity (TFP) using a Cobb–Douglas production function. Corporate resilience (*rir*) is measured as the difference between current-period and previous-period TFP[13,14]. The production function is estimated using ordinary least squares (OLS) and takes the following form:

$$lnY_{it} = \alpha lnK_{it} + \beta lnL_{it} + \gamma lnM_{it} + \sigma_{it} \quad (2)$$

Here, Y_{it} denotes total operating revenue of firm i in year t ; K_{it} denotes total assets; L_{it} represents the total number of employees; and M_{it} denotes expenditure on materials and other inputs. TFP for firm i in year t is derived from the regression residual (σ_{it}).

Core explanatory variable. Data factor utilization capability refers to a firm's comprehensive ability to acquire, store, process, analyze, and realize the value of data[1]. This study uses textual analysis to count selected terms in the Management Discussion and Analysis (MD&A) sections of listed firms' annual reports[15]. The natural logarithm of one plus the term count serves as a proxy for data factor utilization capability (*data*).

Control variables. Following previous studies[1,13], this study controls for firm-level characteristics, including firm size (*size*), firm age (*age*), ownership type (*soe*), board size (board), the shareholding percentage of the ten largest shareholders (*top10*), ownership balance (*balance*), and the fixed-asset ratio (*fixed*).

3.3. Data Sources

The sample comprises Chinese A-share listed firms over the period 2011–2024. Firm-level financial and governance data are obtained from the China Stock Market & Accounting Research (CSMAR) database, while the annual-report texts are collected from the official CNINFO website. City-level raw data are drawn from the National Bureau of Statistics of China and the China National Intellectual Property Administration. The sample is processed as follows. First, financial firms are excluded. Second, firms designated as ST and those delisted during the sample period are removed. Third, only observations for firms with at least five consecutive

ISSN: 2692-7608 DOI: 10.6981/FEM.202611_7(11).0006

years without missing data are retained. Fourth, some missing values are supplemented using linear interpolation and related methods. Fifth, continuous variables are winsorized at the 1st and 99th percentiles to mitigate the influence of outliers. Table 1 reports descriptive statistics.

Table 1. Descriptive Statistics

Variable	*Observations*	*Mean*	*Std. dev.*	*Minimum*	*Maximum*
data	33491	2.719	1.412	0.000	7.045
rir	33491	0.068	0.234	-0.690	0.902
size	33491	22.257	1.317	19.997	26.394
age	33491	2.014	0.896	0.000	3.401
soe	33491	0.309	0.462	0.000	1.000
board	33491	2.111	0.196	1.609	2.639
top10	33491	0.586	0.154	0.228	0.905
fixed	33491	0.203	0.149	0.003	0.665
balance	33491	0.987	0.811	0.049	4.075

4. Empirical Results

4.1. Baseline Regression

Table 2. Baseline Regression Results

	(1)	(2)
	rir	*rir*
data	0.014***	0.007***
	(0.002)	(0.002)
size		0.051***
		(0.004)
age		0.097***
		(0.007)
soe		-0.039***
		(0.011)
board		-0.021
		(0.014)
top10		0.320***
		(0.027)
fixed		-0.182***
		(0.023)
balance		0.004
		(0.004)
cons	0.031***	-1.402***
	(0.006)	(0.095)
Controls	No	Yes
Fixed effects	Yes	Yes
N	33491	33491
R^2	0.141	0.162

Note: *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Firm-clustered robust standard errors are reported in parentheses.

ISSN: 2692-7608 DOI: 10.6981/FEM.202611_7(11).0006

Table 2 reports the baseline estimates of the effect of data factor utilization capability on corporate resilience. Column (1) includes only the core explanatory variable and fixed effects; column (2) additionally incorporates the full set of control variables. The estimated coefficient on data factor utilization capability (*data*) is positive and statistically significant at the 1% level in both specifications, indicating that stronger data factor utilization capability significantly improves corporate resilience. These results are consistent with H1. Effective use of data factors enables firms to identify external risks more sensitively and optimize resource allocation, strengthening their ability to adapt and recover under uncertainty and thereby increasing overall resilience.

4.2. Robustness Checks

Five robustness checks are conducted to assess the reliability of the baseline findings. First, the dependent variable is reconstructed: firm-level TFP is re-estimated using the Olley–Pakes (OP) method and then used to calculate an alternative measure of corporate resilience (*rir2*). Second, allowing for a potential lag in the effect of data factor utilization capability on resilience, the dependent variable (*rir*) is lagged by one period. Third, to account for the influence of other external shocks, observations from 2019 to 2021 are excluded. Fourth, the clustering level of standard errors is changed from the firm level to the city in which the firm is located. Fifth, an additional control for the cash-flow ratio (*cashflow*), measured as net cash flow from operating activities divided by year-end current liabilities, is included in the baseline model. As shown in Table 3, the estimates remain consistent with the baseline results across these checks, supporting their robustness.

Table 3. Robustness Check Results

	Recalculated dependent variable	*Lagged dependent variable*	*Restricted sample period*	*Alternative clustering level*	*Additional control*
	(1)	(2)	(3)	(4)	(5)
data	0.005**	0.009***	0.008***	0.007***	0.007***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
cons	-0.974***	-2.286***	-1.344***	-1.395***	-1.429***
	(0.096)	(0.109)	(0.105)	(0.098)	(0.095)
Controls	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes
N	33491	33491	24864	33491	33491
R^2	0.152	0.162	0.169	0.162	0.169

Note:, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Cluster-robust standard errors are reported in parentheses. Standard errors are clustered at the firm level in columns (1)–(3) and (5), and at the city level in column (4).*

4.3. Endogeneity Test

To alleviate potential endogeneity concerns, this study employs an instrumental-variable approach. The instrument (*IV*) is defined as the mean data factor utilization capability (*data*) of other firms within the same industry–region group. Two-stage least squares (2SLS) estimation is then performed[16], with the results presented in Table 4. In the first stage, the coefficient on the instrumental variable is positive and significant at the 1% level, consistent with theoretical expectations. In the second stage, the coefficient on data factor utilization capability (*data*) is positive and significant at the 10% level. Moreover, the Cragg–Donald Wald F statistic is 147.900, substantially exceeding the rule-of-thumb threshold of 10, which suggests that weak-

ISSN: 2692-7608 DOI: 10.6981/FEM.202611_7(11).0006

instrument concerns are limited. Thus, the baseline conclusion remains unchanged after addressing potential endogeneity.

Table 4. Instrumental-Variable Regression Results

	(1)	(2)
	data	*rir*
IV	0.111***	
	(0.014)	
data		0.047*
		(0.027)
cons	-2.758***	
	(0.626)	
Cragg–Donald Wald F statistic	147.900	
Controls	Yes	Yes
Fixed effects	Yes	Yes
N	33491	33491
R^2	0.804	0.006
Note: Significance notation is the same as in Table 2.		

5. Mechanism Analysis

5.1. Mediation Model Specification

To further test the mechanisms proposed in the theoretical analysis, the following mediation models are specified:

$$M_{i,t} = \beta_0 + \beta_1\, data_{i,t} + \gamma X_{i,t} + \delta_i + \lambda_t + \varepsilon_{i,t} \tag{3}$$

$$rir_{i,t} = \theta_0 + \theta_1\, data_{i,t} + \theta_2\, M_{i,t} + \gamma X_{i,t} + \delta_i + \lambda_t + \varepsilon_{i,t} \tag{4}$$

Here, M denotes the mediating variable linking data factor utilization capability to corporate resilience; the remaining variables are defined as in the baseline model. The coefficient β_1 captures the effect of data factor utilization capability on the mediator, and θ_2 captures the estimated effect of the mediator. Statistically significant coefficients β_1 and θ_2 are interpreted as evidence consistent with mediation.

5.2. Tests of the Proposed Mechanisms

5.2.1. Internal Control.

Internal control is measured using the DIB Internal Control Index (IC)[1]. Columns (1) and (2) of Table 5 present the mechanism test results for internal control. Column (1) shows that the coefficient on data factor utilization capability (data) is significantly positive at the 1% level, indicating that stronger application capability improves internal control quality. In conjunction with Column (2), the results suggest that improvements in firms' data factor utilization capability enhance internal control quality, helping firms reduce decision-making errors and resource losses during external shocks. This strengthened internal control further improves risk identification, loss containment, and operational recovery, ultimately enhancing corporate resilience.

ISSN: 2692-7608 DOI: 10.6981/FEM.202611_7(11).0006

5.2.2. Persistent Innovation Capability.

Persistent innovation capability is measured using a ratio constructed from current and previous R&D expenditure[8]. The specific formula is as follows:

$$rds_{it}=\ln\left[\frac{rd_{it}+rd_{i,t-1}}{rd_{i,t-1}+rd_{i,t-2}}\times\left(rd_{it}+rd_{i,t-1}\right)+1\right] \quad (5)$$

Columns (3) and (4) of Table 5 report the mechanism tests for persistent innovation. In column (3), the coefficient on data is positive and significant at the 5% level. This indicates that stronger data factor utilization capability significantly promotes persistent innovation. Together with column (4), the results suggest that stronger data factor utilization capability promotes knowledge accumulation and technological iteration. This supports persistent innovation and strengthens technological renewal and operational recovery during shocks, ultimately improving corporate resilience.

Table 5. Mechanism Test Results

	(1)	(2)	(3)	(4)
	IC	*rir*	*rds*	*rir*
data	3.590***	0.005**	0.021**	0.007***
	(0.840)	(0.002)	(0.010)	(0.002)
IC		0.001***		
		(0.000)		
rds				0.016***
				(0.003)
cons	323.323***	-1.580***	-9.670***	-0.926***
	(46.236)	(0.096)	(0.629)	(0.118)
Controls	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes
N	33491	33491	33491	33491
R^2	0.286	0.200	0.827	0.201
Note: Significance notation is the same as in Table 2.				

6. Heterogeneity Analysis

6.1. Industry Heterogeneity

Firms in different industries differ in their dependence on data, technology, and knowledge. To examine this heterogeneity, the sample is divided into firms in high-tech and non-high-tech industries. The results in Table 6 show that the estimated coefficient on data factor utilization capability (*data*) is positive and significant at the 1% level for the high-tech subsample but insignificant for the non-high-tech subsample. This indicates that data factor utilization capability plays a more pronounced role in enhancing resilience among high-tech firms. A possible explanation is that high-tech industries are characterized by substantial R&D investment, rapid technological iteration, and pronounced operational uncertainty. Their stronger reliance on high-quality information and knowledge allows the risk-identification and innovation-enhancing benefits of data applications to be realized more readily.

ISSN: 2692-7608 DOI: 10.6981/FEM.202611_7(11).0006

6.2. Heterogeneity Across Innovation Environments

The regional innovation environment may shape firms' ability to absorb data and convert them into economic value. This study uses the logarithm of the total number of patent applications filed in the firm's city during the relevant year to proxy for the local innovation environment. The sample is then split into groups with more and less favorable innovation environments using the mean value of this indicator. As shown in Table 6, the estimated coefficient on data factor utilization capability (*data*) is positive and significant at the 1% level in firms with more favorable innovation environments but insignificant in firms with less favorable environments. These results indicate that the resilience-enhancing effect of data factor utilization capability is more pronounced when the surrounding innovation environment is stronger. A possible explanation is that supportive innovation environments help firms access external knowledge and technological support more quickly, enabling them to adjust business strategies and identify new development opportunities following external shocks.

Table 6. Heterogeneity Test Results

	Industry heterogeneity	*Industry heterogeneity*	*Innovation environment*	*Innovation environment*
	High-tech industries	*Non-high-tech industries*	*More favorable*	*Less favorable*
data	0.009***	0.005	0.008***	0.003
	(0.003)	(0.003)	(0.003)	(0.003)
cons	-1.312***	-1.326***	-1.574***	-1.098***
	(0.124)	(0.155)	(0.139)	(0.143)
Controls	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes
N	20734	12757	19196	14295
R^2	0.093	0.073	0.077	0.092
Note: Significance notation is the same as in Table 2.				

7. Conclusion and Policy Implications

Using data on Chinese A-share listed firms from 2011 to 2024, this study examines the relationship between data factor utilization capability and corporate resilience. Three main conclusions are as follows: First, stronger data factor utilization capability significantly enhances corporate resilience. Second, internal control and persistent innovation constitute two important mechanisms through which this capability improves resilience. Third, the positive effect is more pronounced among firms in high-tech industries and those operating in more favorable external innovation environments.

These findings suggest several policy priorities. First, policymakers could improve the institutional framework governing data as a factor of production and strengthen support for firms' data applications. Clearer rules on data rights, sharing and trading, benefit allocation, and security governance could reduce barriers to data use and associated compliance costs. Second, firms could coordinate data use for internal control and persistent innovation. Integrating operational, financial, and supply-chain data could improve monitoring and early risk detection. Market, technology, and R&D data could help firms identify innovation opportunities and allocate resources. Sharing data and knowledge across departments and supply-chain partners could also support continuity in innovation. Third, support measures could reflect differences in firms' industries and regional innovation environments. For non-high-tech firms, capability

ISSN: 2692-7608 DOI: 10.6981/FEM.202611_7(11).0006

building could focus on R&D, supply-chain management, and risk forecasting. In regions with weaker innovation environments, public innovation platforms and knowledge-exchange networks could broaden firms' access to data-related opportunities.

References

[1] Wu, Q., Cao, C., & Wang, Y. (2026). Firms' data factor utilization capability and supply chain resilience: Differential mechanisms of breadth and depth. China Industrial Economics, (3), 170–193. https://doi.org/10.19581/j.cnki.ciejournal.2026.03.008. (In Chinese)

[2] Bai, Y., Li, J., & Wang, Z. (2022). Data factors: Characteristics, mechanisms, and high-quality development. E-Government, (6), 23–36. https://doi.org/10.16582/j.cnki.dzzw.2022.06.002. (In Chinese)

[3] Xie, K., Xia, Z., & Xiao, J. (2020). The enterprise realization mechanism of big data becoming a real production factor: From the product innovation perspective. China Industrial Economics, (5), 42–60. https://doi.org/10.19581/j.cnki.ciejournal.2020.05.014. (In Chinese)

[4] Nagaraj, A. (2022). The private impact of public data: Landsat satellite maps increased gold discoveries and encouraged entry. Management Science, 68(1), 564–582. https://doi.org/10.1287/mnsc.2020.3896

[5] Hu, H., Song, X., & Dou, B. (2022). The value of digitalization in times of crisis: Evidence of corporate resilience. Finance & Trade Economics, 43(7), 134–148. https://doi.org/10.19795/j.cnki.cn11-1166/f.20220706.008. (In Chinese)

[6] Jiang, L., Ling, Y., Zhang, J., et al. (2022). How does digital transformation affect firm's resilience? An ambidextrous innovation view. Journal of Technology Economics, 41(1), 1–11. (In Chinese)

[7] Dubey, R., Gunasekaran, A., Childe, S. J., Bryde, D. J., Giannakis, M., Foropon, C., Roubaud, D., & Hazen, B. T. (2020). Big data analytics and artificial intelligence pathway to operational performance under the effects of entrepreneurial orientation and environmental dynamism: A study of manufacturing organisations. International Journal of Production Economics, 226, 107599. https://doi.org/10.1016/j.ijpe.2020.107599

[8] Yang, R., & Yang, M. (2025). Persistent innovation effect of digital transformation. Journal of Quantitative & Technological Economics, 42(2), 109–129. https://doi.org/10.13653/j.cnki.jqte.2025.02.001. (In Chinese)

[9] Wang, X., & Shi, Y. (2025). Data elements sharing and corporate resilience: Empirical evidence from public data openness. Contemporary Finance & Economics, (8), 84–98. https://doi.org/10.13676/j.cnki.cn36-1030/f.20250227.002. (In Chinese)

[10] Gu, C., Zhang, S., Zhang, P., et al. (2025). Data empowerment: Market-based allocation of data elements and shaping enterprise resilience. China Soft Science, (12), 166–176. (In Chinese)

[11] Doyle, T. J., Ge, W., & McVay, S. (2007). Accruals quality and internal control over financial reporting. The Accounting Review, 82(5), 1141–1170. https://doi.org/10.2308/accr.2007.82.5.1141

[12] Ghasemaghaei, M., & Calic, G. (2020). Assessing the impact of big data on firm innovation performance: Big data is not always better data. Journal of Business Research, 108, 147–162. https://doi.org/10.1016/j.jbusres.2019.10.029

[13] Tian, D., Ding, B., & Liu, F. (2025). "Patient capital" enabling resilience of startups: The differentiated roles of corporate venture capital and independent venture capital. China Industrial Economics, (7), 156–173. https://doi.org/10.19581/j.cnki.ciejournal.2025.07.008. (In Chinese)

[14] Lu, X., Chiang, Y. C., Wu, F., et al. (2025). Whose AI matters? Examining the bilateral effects of AI capability orientation on supply chain resilience. International Journal of Operations & Production Management, 45(8), 1549–1577. https://doi.org/10.1108/IJOPM-02-2024-0098

[15] Shi, Q., Niu, Y., & Xu, H. (2023). Research on the mechanism of the impact of enterprise data element utilization level on investment efficiency: Using data elements to activate the intermediary role of redundant resources. Journal of Central University of Finance & Economics, (11), 105–115. https://doi.org/10.19681/j.cnki.jcufe.2023.11.009. (In Chinese)

ISSN: 2692-7608 DOI: 10.6981/FEM.202611_7(11).0006

[16] Zhang, J., & Fan, Y. (2024). Innovation input and enterprise resilience: Internal mechanism and industrial chain cooperation. Business and Management Journal, 46(5), 51–71. https://doi.org/10.19616/j.cnki.bmj.2024.05.003. (In Chinese)