Research on the Impact of Logistics Standardization on the Utilization of Enterprise Data Elements

Yi Zheng[1]*

[1] School of International Economics and Trade, Hunan Sany Polytechnic College, Changsha, Hunan, China

Abstract

Against the backdrop of the deepening of the "data element" action, logistics standardization, as an important carrier for the integration of data elements and the real economy, has not been systematically tested for its impact on the utilization of enterprise data elements. Using annual observations of Chinese A-share enterprises as samples, the logistics standardization pilot policy is regarded as a quasi natural experiment, and a multi period double difference model is used for empirical testing. Research has found that logistics standardization significantly promotes the improvement of enterprise data element utilization level; Customer concentration, supplier concentration, and market power all play a partial intermediary role; The relationship between R&D investment intensity and negative regulation; The promotion effect mainly exists in non service industry enterprises and has no significant impact on service industry enterprises. The research enriches the relevant interdisciplinary studies and provides policy references for releasing the value of data elements based on circulation standardization.

Kewords

logistics standardization; Utilization of data elements; Multi period double difference; supply chain structure

1. Introduction

This article takes Chinese A-share listed companies as research samples, and uses the logistics standardization pilot policy as a quasi natural experiment to construct a multi period difference in differences (DID) model to systematically examine the impact of logistics standardization on the utilization of enterprise data elements. On this basis, we will examine the partial mediating effects of customer concentration and market power, examine the negative moderating effect of R&D investment intensity, and conduct heterogeneity analysis by distinguishing between non service and service industries. This article conducts robustness tests through methods such as parallel trend testing, variable substitution, and shortening windows to alleviate endogeneity issues as much as possible and ensure the reliability of conclusions. The marginal contribution of this article is reflected in three aspects: firstly, for the first time, logistics standardization and enterprise data element utilization are included in a unified analysis framework, and the causal effects of the two are empirically identified, expanding the cross research of circulation standardization and data element value; Secondly, reveal the triple intermediary mechanism of "supply chain structure optimization+market power enhancement", clarify the policy transmission path, and

enrich theoretical explanations; Thirdly, identify the moderating effects of R&D investment and industrial heterogeneity, and find that policy dividends are more concentrated in non service industries and low R&D investment enterprises, providing empirical evidence for differentiated policy implementation.

2. Research Design

（1） Sample selection and data sources

This article takes Chinese A-share listed companies from 2010 to 2024 as the initial research sample, and follows the general practice of logistics standardization related research to screen the initial sample as follows: exclude financial listed companies; Exclude listed companies with ST, * ST, and delisting status; Eliminate observations with missing key variable data; Perform bilateral truncation at the 1% level on all continuous variables to eliminate the interference of extreme values on the regression results. Finally, 31138 valid annual observations were obtained from enterprises.

（2） variable definition

Dependent variable: Comprehensive index of enterprise data elements (Data), constructed using the arithmetic mean method of "macro level data element overall index+enterprise level data element development and utilization index" based on existing research (Dai Kuizao et al., 2024). The specific steps are as follows: First, construct the macro level data element overall index (Sbigdatapt). Using the Big Data Development Index of 31 provinces, cities, and districts evaluated in the "Big Data Blue Book: China Big Data Development Report" from 2016 to 2024 as the basic data, and referring to the research methods of previous scholars (Zhao Yunhui et al., 2019) to calculate the average annual growth rate (GP) at the provincial level, the data from 2011 to 2024 is calculated. The specific calculation formula is as follows:

$$Sbigdata_{pt} = Sbigdata_{p,t} \times (1 + g_p)^{2024-t} (t = 2011,2012,\ldots,2024) \quad (1)$$

Among them, p represents the province, t represents the year, Sbigdatap, t is the big data development index of province p in the t-th year, and is the average annual growth rate of the province's big data development index. The second step is to construct an Inddatait index for the development and utilization of data elements at the enterprise level. Based on policy documents and research reports such as the "14th Five Year Plan" for the development of the digital economy and the "White Paper on Data Elements", relevant keywords were selected from four dimensions: data element stock, data development capability, data-driven commercial applications, and data value monetization. Through Python word segmentation processing, manual recognition, and keyword search matching, negative expressions were removed, and the effective keyword total word frequency (Countit) was summarized. The processing formula is as follows:

$$Inddata_{it} = ln(Count_{it} + 1) \quad (2)$$

Among them, i represents the enterprise, Countit is the total frequency of effective keywords for the enterprise in the year, and after adding 1, the natural logarithm is taken to alleviate the problem of data skewness. The third step is to construct a comprehensive index of enterprise level data elements (Datait). Firstly, standardize Sbigdatapt and Inddatait to eliminate

dimensional differences, and then use the arithmetic mean method to calculate the composite index. The specific formula is as follows:

$$Data_{it} = \frac{Std(Sbigdata_{pt}) + Std(Inddata_{it})}{2} \quad (3)$$

$$Std(X) = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (4)$$

Among them, Std (X) represents the standardization processing function, which converts the indicator value into a standardized value within the [0,1] interval; Sbigdatapt matches the annual data of the corresponding province based on the registered location of the enterprise. The larger the Datait value, the higher the level of utilization of enterprise data elements.

Core explanatory variable: Logistics Standardization Policy (DID), which is a multi period double difference interaction term. If the city where the enterprise is registered becomes a pilot city for logistics standardization in the current year or later, the value is 1; otherwise, it is 0.

Control variables: Referring to existing research, select the following control variables that affect the utilization of enterprise data elements: profitability, measured by return on equity; Enterprise growth, measured by the rate of revenue growth; Enterprise size, measured by taking the natural logarithm of total assets; The shareholding ratio of the largest shareholder (top 1) is measured by the proportion of the number of shares held by the largest shareholder to the total number of shares; Cash flow status, measured by net cash flows from operating activities; The proportion of the tertiary industry in a region (indus3) is measured by the ratio of the added value of the tertiary industry in the city where the enterprise is located to the GDP.

Mediating variable: Customer concentration (CUS5), measured by the proportion of sales revenue from the previous five major customers to the total sales revenue; Supplier concentration (Sup5), measured by the proportion of the previous five major suppliers' procurement amount to the total procurement amount; Market power (Lerner), measured by the Lerner index.

Adjusting variable: R&D investment intensity (rdint), measured as the ratio of R&D expenditure to operating revenue.

Table 1 Descriptive statistical results of each variable

variable	Sample value	average	standard deviation	minimum	maximum
data	31138	3.959	1.282	0.000	7.622
DID	31138	0.384	0.486	0.000	1.000
profit	31138	0.030	0.086	-3.994	0.786
growth	31138	0.134	0.491	-0.928	45.460
ln_size	31138	22.282	1.303	17.806	28.791
top1	31138	33.190	14.717	0.000	89.990
cash_net	31138	1.675	31.593	-1640.052	4887.906
indus3	31138	57.638	13.088	10.150	85.270
rd_int	31138	5.334	6.504	-73.410	227.710
cus5	31138	33.143	22.253	-34.600	106.620
lerner	31138	0.113	0.173	-0.981	2.203

(3) Model construction

Baseline regression model; To examine the impact of logistics standardization on the utilization of enterprise data elements, this

paper considers the pilot work of logistics standardization as a quasi natural experiment and constructs a multi period double difference (DID) model:

$$data_{it} = \beta_0 + \beta_1 DID_{it} + \sum_{k=1}^{n} \gamma_k Contr$$

(5)

Among them, i represents the enterprise and t represents the year; Datait is the level of data element utilization of the enterprise in the year; DIDit is the core explanatory variable; Controlskit is the th control variable; μ i is the fixed effect of enterprises, used to control the heterogeneity characteristics of enterprises that do not change over time; λ t is a fixed year effect used to control for time shocks faced by all enterprises; ε it is a random perturbation term. All regression standard errors in this article are clustered at the enterprise level.

Moderation effect model; To test the moderating effect of R&D investment intensity, an interaction term between R&D investment intensity and the core explanatory variable is added to the benchmark model:

$$data_{it} = \varphi_0 + \varphi_1 DID_{it} + \varphi_2 rd_int_{it} + \varphi_3$$

(6)

Among them, DIDit × rdintit is the interaction term. If ϕ 3 is significantly positive, it indicates that R&D investment intensity positively regulates the impact of logistics standardization on the utilization of enterprise data elements; If it is significantly negative, it is a negative adjustment.

Mediation effect model; Using stepwise regression to test the mediating effect, construct the following three models in sequence:

$$data_{it} = \alpha_0 + \alpha_1 DID_{it} + \sum_{k=1}^{n} \alpha_k Contr$$

(7)

$$M_{it} = \delta_0 + \delta_1 DID_{it} + \sum_{k=1}^{n} \delta_k Controls_{kit} + \mu_i + \lambda_t + \varepsilon$$

(8)

$$data_{it} = \theta_0 + \theta_1 DID_{it} + \theta_2 M_{it} + \sum_{k=1}^{n} \theta_k Controls_{kit} +$$

(9)

Among them, are Mit mediator variables (cus5, sup5, lerner). If α 1, δ 1, and θ 2 are all significant, and θ 1 is significantly smaller than α 1, then there is a partial mediating effect; If α 1, δ 1, and θ 2 are all significant, and θ 1 is not significant, then there is a complete mediating effect.

3. Empirical analysis

(1) Multicollinearity test

The variance inflation factor (VIF) test was conducted, and the results showed that the VIF values of all variables were far less than 10, with a maximum value of 1.44 and a mean of about 1.1, confirming that the model did not have serious multicollinearity issues and the regression results were reliable.

Table 2: Multicollinearity Test

variable	VIF	1/VIF
DID	1.42	0.70
profit	1.06	0.95
growth	1.03	0.97
ln_size	1.05	0.95
top1	1.06	0.94
cash_net	1.00	1.00
indus3	1.44	0.70

(2) Benchmark regression analysis

This article constructs a multi period double difference model to test the impact of logistics standardization on the utilization of enterprise data elements. The robustness of the results is verified by gradually adding

control variables. All models control for individual fixed effects and year fixed effects of enterprises, and the standard errors are clustered at the enterprise level. The results showed that as the control variables were gradually added, the coefficient of the core explanatory variable DID slightly decreased from 0.088 to 0.054, but remained significantly positive at the 1% statistical level, indicating that the promotion effect of logistics standardization on the utilization of enterprise data elements was not affected by the selection of control variables, and the main effect was robust. The complete model in column (7) indicates that, while controlling for other factors, the logistics standardization pilot policy can increase the data element utilization level of enterprises in pilot cities by an average of 5.4%. In terms of controlling variables, the coefficients of enterprise size, revenue growth rate, and industry digitalization degree are significantly positive, which is in line with economic expectations.

(3) Adjustment effect analysis

The regression results with R&D investment intensity (rdint) as the moderating variable are shown in Table 3. Column (1) is the baseline regression result, and column (2) is the regression result after adding moderating variables and interaction terms. It can be observed that the coefficient of the interaction term is -0.010, which is significantly negative at the 1% level, indicating that R&D investment intensity has a significant negative moderating effect on the relationship between logistics standardization and enterprise data element utilization.

Table 3 Analysis of Regulatory Effects

variable	(1) data	(2) data
DID	0.054***	0.109***
	(0.019)	(0.021)
profit	0.039	0.050
	(0.045)	(0.045)
growth	0.017**	0.017**
	(0.008)	(0.008)
ln_size	0.251***	0.252***
	(0.008)	(0.008)
top1	0.001	0.001
	(0.001)	(0.001)
cash_net	-0.000**	-0.000**
	(0.000)	(0.000)
indus3	0.009***	0.009***
	(0.001)	(0.001)
rd_int		0.006***
		(0.001)
c.DID#c.rd_int		-0.011***
		(0.002)
_cons	-2.201***	-2.263***
	(0.198)	(0.199)
id	Yes	Yes
year	Yes	Yes
N	31138	31138
r2	0.859	0.859
r2_a	0.837	0.838
F	150.753	121.243

$^{*}p<0.1, ^{**}p<0.05, ^{***}p<0.01$

This result indicates that the higher the R&D investment intensity of enterprises, the weaker the promotion effect of logistics standardization on their data element utilization. The possible reason is that high R&D investment enterprises usually have established relatively complete internal data management and application systems, and the incremental effect of data elements brought by logistics standardization is relatively limited; However, the data foundation of low R&D investment enterprises is relatively weak, and logistics standardization can provide them with standardized data interfaces and processes, significantly improving their level of data element utilization.

(4) Analysis of intermediary effect

This article examines the mechanism by which logistics standardization affects the utilization of enterprise data elements from two dimensions: customer concentration and market power.

The benchmark regression results show that the overall effect of DID on the data is significantly positive, and the main effect holds. The coefficient of DID for customer concentration (cus5) is -1.683, which is significantly negative at the 1% statistical level, indicating that logistics standardization can effectively reduce customer concentration in enterprises. After including both DID and cus5, the coefficient of cus5 is -0.002, which is significantly negative at the 1% statistical level, and the coefficient of DID is still significantly positive. This indicates that customer concentration plays a partial intermediary role, and logistics standardization forces enterprises to integrate more customer demand data and transaction data by reducing customer concentration, thereby improving the utilization level of data elements. Assuming H3a is valid.

The coefficient of DID for market power (lerner) is 0.034, which is significantly positive at the 1% statistical level, indicating that logistics standardization can significantly enhance a company's market power. After incorporating both DID and Lerner, the coefficient of Lerner is 0.085, which is significantly positive at the 1% statistical level, and the coefficient of DID remains significantly positive. This indicates that market forces play a partial intermediary role, and logistics standardization enhances the market power of enterprises, increases digital resource investment and cross enterprise data collaboration, thereby promoting the utilization of data elements. Assuming H3b holds true.

In summary, logistics standardization promotes the utilization of enterprise data elements through three paths: "reducing customer concentration - enhancing data element utilization" and "enhancing market power - enhancing data element utilization", all of which are partial intermediary effects.

4. Conclusion

This article takes A-share listed companies from 2010 to 2024 as samples, and uses logistics standardization pilot as a quasi natural experiment to test its impact on the utilization of enterprise data elements using a multi period double difference model. The results show that, firstly, assuming H1 is true, logistics standardization significantly improves the utilization level of enterprise data elements, and the pilot policy can increase the average utilization level of relevant enterprise data elements by 5.4%. Multiple robustness tests have confirmed the causal relationship. Secondly, assuming H2 holds true, R&D investment intensity exhibits a significant negative moderating effect, with high R&D enterprises having a well-established digital foundation and

weak policy marginal effects; Low R&D enterprises can make up for their shortcomings through standardization, and the policy dividends will be more prominent. Thirdly, assuming H3a and H3b are established, customer concentration (intermediary proportion 6.23%) and market power (intermediary proportion 5.35%) play a partial intermediary role, with the total intermediary effect accounting for 11.58%. Fourthly, assuming H4 holds true, there is industrial heterogeneity in policy effects, which is only significant for non service industry enterprises and has limited impact on service industry enterprises.

References

[1]Lu C ,Chen J ,Yang X , et al.How Digital Inclusive Finance Affect the Exports of Small and Medium Enterprises: Based on Growth Enterprise Market Data in China[J].Emerging Markets Finance and Trade,2026,62(14):5753-5767.DOI:10.1080/1540496X.2026.2630751.

[2]The Four Data Engineering Debts That Quietly Undermine Enterprise Data Analytics[J].M2 Presswire,2026,

[3]Fang Y ,Ma Z .Through the lens of transaction costs: The impact of logistics standardization on corporate ESG performance[J].Transport Policy,2026,189104406-104406.DOI:10.1016/J.TRANPOL.2026.104406.

[4]Jin T ,Liang X ,Lu M .Does logistics service standardization matter for digital transformation? Evidence from a natural experiment[J].Applied Economics Letters,2026,33(15):2868-2875.DOI:10.1080/13504851.2025.2490717.

[5]Zhang Q ,Zhao S ,Zhang M .Standardization of industry norms and urban entrepreneurial activity:Empirical evidence from China's logistics standardization construction[J].Economic Analysis and Policy,2026,901089-1104.DOI:10.1016/J.EAP.2026.02.007.

[6]Rui J ,Gaosong Z ,Fei W .LOGISTICS SERVICE STANDARDIZATION, ENTERPRISE OPERATION EFFICIENCY, AND ECONOMIC EFFECTS[J].TRANSFORMATIONS IN BUSINESS & ECONOMICS,2019,18(3):168-190.