The Effect of Data Asset Disclosure on the Innovation Input Level of Listed Companies based on DID Model

Lu Lin *

School of Business Administration of Jimei University, Xiamen, China

* Corresponding Author Email: Lin_Lu_@outlook.com

Abstract. Against the backdrop of the vigorous development of the digital economy, the implementation of the Interim Provisions on the Accounting Treatment of Enterprise Data Resources on January 1, 2024, holds milestone significance. This paper focuses on data asset disclosure in financial statements, selects the annual report data of listed companies in 2023 and 2024 as samples, and uses the difference-in-differences (DID) method to conduct a study on the impact of data asset disclosure in financial statements on the innovation input level of listed companies. The results show that data asset disclosure can significantly improve enterprises' R&D input level, and the robustness test further enhances the reliability of the study. This research outcome not only provides practical guidance for enterprises to use data asset disclosure to enhance their innovation capabilities, but also offers theoretical experience for in-depth understanding of how data factors drive high-quality economic development at the micro level. Meanwhile, it provides empirical references for relevant authorities to optimize the system of data asset disclosure and fully release the value of data factors.

Keywords: Data asset, Innovation, Difference-in-Differences Model.

1. Introduction

Against the backdrop of the global digital economy wave, data, as a new-type factor of production, is increasingly prominent in its value-creating capacity. In recent years, China has prioritized data factor market development and issued key policies. In 2024, data infrastructure accelerated: the National Data Administration and 16 others jointly released the Three-Year Action Plan for "Data Factor ×" (2024-2026) (clarifying goals and 12 initiatives for data factor-driven development); the Ministry of Finance also issued the Guiding Opinions on Strengthening Data Asset Management, defining 12 tasks including compliant data asset management and rights clarification. Against this policy backdrop, the Interim Provisions on the Accounting Treatment of Enterprise Data Resources took effect on Jan 1, 2024. This milestone policy shifts data asset disclosure in financial statements from theory to practice and guarantees market-oriented data factor allocation. Data asset disclosure refers to recognizing and measuring enterprise data resources in financial statements—a reform that changes enterprises' asset structure and may deeply impact their innovation decisions.

From the perspective of existing relevant studies, the concept of "data asset" was first regarded as a portfolio of government and commercial bonds [1]. However, with the rapid iteration of information technology and Internet technology, the connotation of "data asset" has significantly diverged from its original definition. In the context of the big data era, data should be regarded as an asset, and enterprises need to place data in the important category of corporate assets for planning and management during their operation and management processes [2]. The Practice White Paper on Data Asset Management (Version 6.0), released by the Big Data Technology and Standard Committee in 2023, defines data assets as structured or unstructured data stored in electronic form, legally owned or controlled by entities such as government agencies, enterprises and public institutions, and capable of creating economic value through transactions [3]. At the micro level, innovation has gradually become a topic of great concern in the field of data assets. The concept of dynamic capabilities points out that companies have the ability to integrate, build and reconfigure internal capabilities to respond to changes in the business environment [4]. This theory provides a theoretical basis for understanding the dynamic adjustments of enterprises in the process of data assetization. Enterprises must consider building capabilities in data assetization to cope with the ever-changing digital economic environment

[5]. This means that in the process of adapting to the development and changes in the field of data assets, enterprises need to continuously adjust their internal resources and capabilities based on the dynamic capability theory. From an empirical perspective, as a new-type factor of production, data assets are conducive to increasing enterprise innovation input from the perspectives of innovation willingness and innovation resources [6]; data asset information disclosure is beneficial to promoting digital technology innovation, thereby driving the high-quality development of enterprises [7]. However, there are still relatively few studies on the impact of data asset disclosure on enterprises' innovation input level after the formal implementation of the data asset disclosure policy.

This paper aims to explore the impact of data asset disclosure on enterprises' innovation input level. To achieve this goal, this paper selects the annual report data of listed companies in 2023 and 2024 as research samples and uses the difference-in-differences (DID) method for empirical analysis. By constructing a scientific and reasonable econometric model, this paper analyzes the mechanism of action of data asset disclosure on enterprises' innovation input level from the micro level, hoping to provide references for enterprises to deploy data assets and for policies to improve the data factor market, promote data assets to deeply empower enterprise innovation, and help the high-quality development of the economy reach a new stage.

2. Theoretical Analysis and Hypothesis Development

In the view of traditional economics, the sole operating objective of an enterprise is to maximize corporate value [8]. As a key way for enterprises to break through resource constraints and increase product added value, innovation is one of the main means for enterprises to obtain excess economic profits. However, due to the characteristics of innovation projects such as long cycles and high costs (Holmstrom) [9], managers often focus on short-term profits and tend to choose conventional investments with lower risks, while lacking sufficient incentives for innovation projects that are more risky but can drive the long-term development of enterprises [10].

Data asset disclosure provides a new solution to this problem. Data assetization is a process of activating the potential value of data and promoting its transformation from a virtual form to a value-based form. Its core feature is that data is effectively managed as an asset, with its economic attributes and property rights highlighted. When assetized data enters the economic field to participate in profit distribution and acquires the attribute of value appreciation, it evolves into data capital [11]. This logic of value transformation and appreciation also gives enterprises inherent motivation to promote data assetization. Policies clearly stipulate that enterprise data resources can be recognized and measured in financial statements: the higher the quality of an enterprise’s data resources, the broader their application scenarios, and the stronger their value-creating capacity, the higher the asset value measured in the statements. This accounting rule effectively guides enterprises to attach importance to the accumulation, governance, and in-depth development of data resources, and promotes investment in data-driven R&D technologies. In addition, as a strategic asset of enterprises, data assets can improve operational efficiency and reduce production and operation costs [12]. The explicit valuation of an enterprise’s data resources can optimize its asset structure, enhance its financing capacity, and provide more stable financial support for R&D projects [13]. It can be seen that the data asset disclosure policy effectively promotes the improvement of enterprise’s R&D input level through the mechanisms of explicit data value and optimized resource allocation. Based on this, this paper proposes the following hypothesis to be verified:

Hypothesis: Data asset disclosure in financial statements can promote listed companies to increase their innovation input level.

3. Research Design

3.1. Sample Selection and Data Sources

This study's initial sample covers Chinese A-share listed companies from Jan 1, 2023, to Dec 31, 2024, screened by: (1) Treatment group: 100 firms that recognized data assets in 2024 annual reports; control group: 1 same-industry non-recognizing A-share firm per treatment firm (total 100); (2) Excluding ST/*ST firms and those with abnormal financial data; (3) Removing samples with missing values; (4) 1% Winsorization for all continuous variables (trimming top/bottom 1%) to avoid extreme value interference.

Data on data asset disclosure and listed firms' financials mainly comes from Tracking Report on Data Asset Recognition in Financial Statements of Chinese Enterprises (Shanghai Advanced Institute of Finance, Shanghai Jiao Tong University), RESSET Database, and firms' annual reports.

3.2. Model Design

The Difference-in-Differences (DID) method is adopted to evaluate the impact of data asset disclosure in financial statements on the level of innovation input. The following baseline Model (1) is constructed:

$$\mathrm{Rd\_spend}_{it} = \beta_0 + \beta_1 Post_t + \beta_2 Treat_i + \beta_3 (Treat_{it} \times Post_{it}) + \beta_4 Controls_{it} + \mu_i + \lambda_i + \varepsilon_{it} \quad (1)$$

$\beta_1$represents the average overall impact on the subjects after the policy implementation;$\beta_2$represents the baseline level of the treatment group before the policy implementation; $\beta_3$is the key coefficient, indicating the actual impact of the policy on the treatment group, i.e., the additional change of the treatment group compared with the control group after the policy implementation.

3.3. Variable Selection

3.3.1. Dependent Variable

The dependent variable is enterprise innovation input (R&d). A key criterion for identifying high-tech enterprises, it has strong cross-sectional comparability, enabling effective comparison across differing enterprises. Thus, innovation input is a core factor boosting competitiveness and driving transformation. This study uses annual R&D expenditure as the core indicator.

3.3.2. Independent Variable

The independent variable is data asset disclosure in financial statements (Treat). Following Xue et al. [14], the treatment group (companies disclosing data assets per accounting standards) is coded 1; the control group (1 same-industry non-disclosing A-share listed company) is coded 0. A policy time dummy (Post) is included: 1 for years after, 0 before policy implementation. The core interaction term Treat × Post measures the net effect on R&D input.

While 2024 data disclosure might affect 2025 innovation input (with annual reports usually disclosed by May next year), only 2024 annual report data is available since the 2024 Interim Provisions. As annual report disclosure decisions require in-year work, this study treats their effect as impacting the same year—i.e., the influence of current-year data asset disclosure decisions on innovation input.

3.3.3. Control Variables

Following Kun Qian et al. [15] and considering enterprise R&D input drivers, control variables include enterprise size (Size), listing age (Age), return on total assets (Roa), R&D personnel proportion (Rd), and asset-liability ratio (Lev). These variables control for enterprise traits (scale, tenure, profitability, R&D workforce, solvency) that may affect R&D input. Specific variables and calculations are in Table 1.

Table 1. Definition of Main Variables

Variable Category	Variable Name	Variable Symbol	Calculation Method
Dependent Variable	R&D Input	R&d	Annual R&D expenditure of the company
Core Independent Variables	Data Asset Disclosure Group	Treat_i	1 = Treatment group (companies with data asset disclosure); 0 = Control group (companies without data asset disclosure)
	Policy Time Dummy Variable	Post_t	1 = After the implementation of the data asset disclosure policy; 0 = Before the implementation
	DID Interaction Term	Treat × Post	Product of Data Asset Disclosure Group and Policy Time Dummy Variable
Control Variables	Enterprise Size	Size	Natural logarithm of total assets
	Listing Age	Age	Difference between the current year and the listing year
	Return on Total Assets	Roa	Ratio of total profit to average total assets
	Proportion of R&D Personnel	Rd	Ratio of the number of R&D personnel to the total number of employees
	Asset-Liability Ratio	Lev	Ratio of total liabilities to total assets

4. Empirical Results Analysis

4.1. Descriptive Statistics

Table 2 shows descriptive statistics for key variables. Enterprise R&D input (R&d) has a standard deviation of 4.03×10^{9}, with min 9.61×10^{5} and max 2.67×10^{10}, indicating significant differences across enterprises. Post-policy period (Post) has a mean of 0.467 (46.7% of observations post-implementation). Treatment group (Treat) mean is 0.530 (53.0% of sample), control group 47.0%. Interaction term (Treat_Post) mean 0.261 (26.1% of sample in treatment group post-policy).

Table 2. Results of descriptive statistics

VarName	Obs	Mean	SD	Min	Max
R&d	327	1.26e+09	4.03e+09	9.61e+05	2.67e+10
Post	327	0.468	0.500	0.000	1.000
Treat	327	0.532	0.500	0.000	1.000
Treat_Post	327	0.263	0.441	0.000	1.000
Size	327	23.512	1.694	20.413	28.251
Age	327	14.364	9.041	1.000	31.000
ROA	327	0.017	0.058	-0.190	0.130
Rd	327	17.541	17.361	0	63.710
Lev	327	0.475	0.197	0.074	0.860

4.2. Baseline Regression

This paper conducts regression analysis on Model (1). Column (1) of Table 3 presents the regression results with only firm fixed effects and year fixed effects controlled. The regression coefficient of Treat_Post is 1.003e+08, which is significantly positive at the 5% significance level. This supports the hypothesis that data asset disclosure can significantly promote listed companies to increase their innovation input level.

This paper further adds control variables in Column (2) of Table 3, and finds that Treat_Post remains significantly positive at the 5% significance level. The regression results in Table 3 support the hypothesis of this paper, indicating that data asset disclosure enhances the motivation of listed companies to increase innovation input, thereby promoting the innovative development of enterprises.

Table 3. Baseline regression results

	(1)	(2)
VARIABLES	R&d	R&d
Post	-5.319e+07	-9.146e+07
	(-1.52)	(-0.51)
Treat	2.735e+08	1.534e+08
	(1.35)	(0.03)
Treat_Post	1.003e+08**	9.873e+07**
	(2.13)	(2.04)
Constant	9.531e+07	-4.772e+09
	(0.66)	(-1.13)
Observations	327	327
R-squared	0.999	0.999
Controls	NO	YES
Firm FE	YES	YES
Year FE	YES	YES

t-statistics in parentheses, *** p<0.01, ** p<0.05, * p<0.1

4.3. Robustness Tests

Since the Interim Provisions on the Accounting Treatment of Enterprise Data Resources took effect on Jan 1, 2024, and only 2024 annual report data is currently available, this paper uses t-tests for each control variable to temporarily replace the parallel trend test.

Table 4 shows core covariate balance test results: no significant differences in core covariates between the treatment and control groups. This suggests good sample balance pre-policy, supporting the regression results' robustness.

Table 4. Balance test of key covariates results

	M±SD Treat=0 (n=155)	M±SD Treat=1 (n=175)	t	p
Size	23.504±1.603	23.523±1.763	-0.106	0.916
Age	15.387±8.553	13.669±9.477	1.721	0.086
ROA	0.018±0.065	0.016±0.051	0.350	0.727
Lev	18.477±16.079	16.459±14.669	1.193	0.234
Size	0.467±0.198	0.483±0.198	-0.706	0.481
Age	23.504±1.603	23.523±1.763	-0.106	0.916

4.4. Heterogeneity Analysis

4.4.1. Heterogeneity Test Based on Enterprise Characteristics

Enterprise nature differences may affect how data assets influence innovation input. Following Li et al. [6], samples are split into state-owned enterprises (SOEs) and non-SOEs by ownership, with separate regressions. Table5 reports ownership-based heterogeneity results: Column (1) for non-SOEs, Column (2) for SOEs.

Results show policy shocks significantly boost non-SOEs' R&D input, but SOEs fail the significance test (no notable impact). This gap may come from differing governance and goals: non-

SOEs depend more on innovation for market advantages and are more policy-sensitive, while SOEs face multi-goal constraints, making their R&D decisions less policy-sensitive.

4.4.2. Heterogeneity Test Based on Industry Nature

Industry differences affect innovation demand. Following Lu and Dang [16], samples are classified into technology-intensive (electronics, pharmaceuticals, IT, machinery/equipment/instrumentation) and non-technology-intensive enterprises, with separate regressions.

Table 5 reports industry-based heterogeneity results: Column (3) for technology-intensive, Column (4) for non-technology-intensive. Policy shocks significantly boost R&D input in technology-intensive enterprises via mechanisms like data assets—linked to their focus on technological innovation as core competitiveness, making them more sensitive to policies/data assets and enabling quick conversion of support into R&D.

Non-technology-intensive enterprises fail the significance test (no notable effect). This may reflect their reliance on traditional operations, where technology is not key to competition, with limited data asset applications and innovation motivation weakening policy response.

Table 5. Results of Heterogeneity Tests

	(1)	(2)	(3)	(4)
VARIABLES	R&d	R&d	R&d	R&d
Post	79271281.934	-1.493e+08	-3.269e+08	-7.379e+07
	(0.49)	(-0.38)	(-1.30)	(-0.32)
Treat	-3.591e+09	3.722e+08	-6.872e+09	-9.809e+08
	(-0.80)	(0.04)	(-1.25)	(-0.15)
Treat_Post	1.204e+08**	9.163e+07	1.838e+08**	7.063e+7
	(2.26)	(1.09)	(2.26)	(1.19)
Constant	-4.632e+09	-1.074e+09	-4.384e+08	-7.657e+09
	(-1.33)	(-0.12)	(-0.05)	(-1.44)
Observations	122	183	100	227
R-squared	0.998	0.999	0.999	0.999
Controls	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

t-statistics in parentheses, *** p<0.01, ** p<0.05, * p<0.1

5. Conclusion

Based on the difference-in-differences (DID) model, this study conducts an in-depth exploration of the impact of data asset disclosure in financial statements on enterprise innovation input. The main research conclusions are as follows: Data asset disclosure in financial statements significantly promotes the increase in enterprise innovation input, which verifies the positive driving effect of data asset disclosure on enterprise innovation input.

Furthermore, from the perspective of enterprise attributes, the innovation incentive effect of data asset disclosure is more prominent in the sample of non-state-owned enterprises. From the perspective of industrial characteristics, the sample of technology-intensive enterprises shows higher significance, reflecting that such enterprises have stronger demand for the application of data assets and higher transformation efficiency.

The research conclusions of this paper indicate that data asset disclosure in financial statements can provide a new path for the innovative development of enterprises. For enterprises, promoting the compliant disclosure of data assets can be transformed into a "power source" for innovation input, helping to break through the bottleneck of R&D resources. For the financial market, the enterprise operation information carried by data assets can alleviate information asymmetry between banks and

enterprises, optimize the allocation of credit resources, and indirectly create a favorable financing environment for enterprise innovation.

However, limited by the short time of data accumulation, it is difficult to fully observe the long-term dynamic effects. In addition, this study does not further explore the impact of data asset disclosure on innovation output or innovation quality. Therefore, in future research, the observation period can be further extended to continuously track the iteration of the data asset disclosure policy and dynamically observe changes in effects. At the same time, indicators such as the number of patents and transformation efficiency can be included to construct a multi-dimensional innovation measurement system.

References

[1] Peterson, Richard E. A cross section study of the demand for money: the United States, 1960-62. The Journal of Finance, 1974, 29.1: 73-88.

[2] Fisher T, Booksx I. The data asset: how smart companies govern their data for business success.Wiley Publishing, 2009.

[3] Big Data Technology and Standard Committee. Practice white paper on data asset management (Version 6.0) [R]. Beijing: Big Data Technology and Standard Committee, 2023.

[4] Peteraf M A. The cornerstones of competitive advantage: A resource-based view [J]. Strategic Management Journal, 1993, 14 (3):179-191.

[5] Chanias S. Mastering Digital Transformation: The Path of a Financial Services Provider towards a Digital Transformation Strategy [C]//European Conference on Information Systems.2017.

[6] Li Jian,Dong Xiaofan,Zhang Jinlin. The impact of data assets on enterprises' innovation input. Foreign Economics & Management, 2023, 45 (12):18-33.

[7] Zhao Xueqi, Sun Xiaozhe, Li Jizhen.The impact of data asset information disclosure on digital technology innovation [J/OL].Studies in Science of Science,2025,1-15 [2025-08-05].

[8] Jensen M C. Value maximization, stakeholder theory, and the corporate objective function [J]. European Financial Management, 2002, 12 (2):235-256.

[9] Holmstrom B. Agency costs and innovation [J]. Journal of Economic Behavior & Organization, 1989, 12.3: 305-327.

[10] KEUM, D. Daniel. Innovation, short-termism, and the cost of strong corporate governance. Strategic Management Journal, 2021, 42.1: 3-29.

[11] SAI, Zhuli; CHENG, Yunian. Difficulties and countermeasures in data asset pricing. In: International Conference on Cloud Computing. Cham: Springer Nature Switzerland, 2023. p. 85-94.

[12] Goodridge P, Haskel J. How does big data affect GDP? Theory and evidence for the UK. jonathan haskel, 2015.

[13] Begenau J, Farboodi M, Veldkamp L. Big data in finance and the growth of large firms. Journal of Monetary Economics, 2018, 97 (AUG.):71-87.

[14] Xue Y., Zhou K., Xu Z..Can Data Asset Recognition in Financial Statements Become an Enterprise Accelerator? — An Effect Measurement of Data Asset Recognition in Financial Statements on Listed Companies' Profits Based on the DID Model. Science Technology and Industry, 2025, 25 (12):212-220.

[15] Qian K, Pan C, Liang X. Does data asset disclosure affect bank lending acquisition? [J]. International Review of Financial Analysis, 2025, 97: 103834.

[16] Lu T., Dang Y. Corporate governance and technological innovation: a cross-industry comparison. Economic Research Journal, 2014, 49 (06):115-128.