Article

Not peer-reviewed version

Can Data Assetisation Boost Corporate Investment Efficiency in the Fintech Context?

Hongying Luo * , Jian Xu , Li Zhu , Yifan Fu

Posted Date: 10 March 2026

doi: 10.20944/preprints202603.0705.v1

Keywords: data assetisation; corporate investment efficiency; financial technology

Article

Can Data Assetisation Boost Corporate Investment Efficiency in the Fintech Context?

Hongying Luo [1,*], Jian Xu [1], Li Zhu [2] and Yifan Fu [3]

[1] School of Economics and Management, University of Chinese Academy of Sciences
[2] School of Finance, Xinjiang University of Finance and Economics, Urumqi
[3] School of Finance, Capital University of Economics and Business
* Correspondence: 2308451099@qq.com

Abstract

Against the backdrop of the high-quality development of the digital economy, exploring the impact of data assetisation on corporate investment efficiency with fintech as the core tool is of great significance for driving the steady and orderly development of enterprises... Taking A-share listed companies from 2012 to 2023 as the research sample, and further classifying them into self-used data assets and transactional data assets based on the types of data assets., the study finds: (1) Data assetisation significantly enhances corporate investment efficiency, with self-use data assets demonstrating a stronger driving effect. (2) Mechanism analysis reveals that data assetisation alleviates underinvestment by easing financing constraints and leveraging the "talent effect". Concurrently, it mitigates overinvestment by reducing agency problems and accelerating digital transformation, thereby enhancing investment efficiency. (3) Heterogeneity tests indicate that the positive impact of data assetisation on investment efficiency is more pronounced among growth-stage enterprises, technology-intensive firms, and companies operating in regions with high bank liquidity. (4) Banking fintech positively moderates the enhancement of corporate investment efficiency through data assetisation, with a more pronounced effect on alleviating underinvestment. However, it may also exacerbate overinvestment. Consequently, enterprises should vigorously develop data assetisation, applying different types of data assets to specific use cases to unlock data dividends. This approach supports the scientific development of corporate investment decisions and enhances investment efficiency.

Keywords: data assetisation; corporate investment efficiency; financial technology

1. Introduction

As one of the traditional "three drivers" of the economy, investment serves as a primary engine for economic growth, playing a crucial role in optimising resource allocation and propelling scientific innovation and technological advancement [1]. The Third Plenary Session of the 20th CPC Central Committee emphasised the need to "deepen reforms in the investment approval system and improve mechanisms to stimulate the vitality of social capital investment and facilitate investment implementation." As of October 2024, national fixed-asset investment reached approximately RMB 42.3222 trillion, representing a year-on-year increase of 3.4%, with investment scale exhibiting a trend of annual growth [1]. However, China's investment efficiency faces dual challenges of regional polarisation and inefficiency, hindering its role in stabilising the broader economic landscape. Entering the "new normal" development phase, ensuring capital markets serve the real economy more efficiently and facilitating a smooth transition from high-speed to high-quality economic growth necessitates prioritising investment efficiency while maintaining reasonable investment scale growth. In recent years, persistent regional trade conflicts and public health crises have heightened

[1] (Citation: Data sourced from the National Bureau of Statistics)

 doi:10.20944/preprints202603.0705.v1

enterprises' perception of risk uncertainty, dampening their investment willingness and efficiency. This has led to a mismatch between the scale of capital input and output. Consequently, effectively boosting corporate investment efficiency to achieve macroeconomic transformation and industrial upgrading has become a critical issue requiring urgent resolution.

Against this backdrop, existing literature examines factors influencing corporate investment efficiency from both external environmental and internal governance perspectives. From an external environment viewpoint, elements such as government governance [2], government quality [3], and public data openness [1] impact corporate investment efficiency by affecting resource allocation efficiency. From an internal governance perspective, factors such as equity structure [4], digital transformation [5], and managerial overconfidence and optimism bias [6] affect corporate investment efficiency by influencing internal agency problems. Some scholars attribute insufficient corporate investment efficiency to underinvestment and overinvestment, identifying financing constraints and internal information asymmetry as key drivers of these issues [7]. In summary, existing literature provides substantial theoretical underpinnings for this study's exploration of factors influencing corporate investment efficiency. However, few scholars have examined the impact of data assetisation on corporate investment efficiency from a data asset perspective and with fintech as the starting point, nor have they explored fintech's pivotal role in this process. This gap presents an opportunity for the present research.

In the digital economy era, the deepening development of fintech provides a broad application platform for data elements to empower corporate production and value creation, while also offering new opportunities for data assets and traditional factors to boost corporate investment efficiency. Data assetisation, as a dynamic process evolving from data → data elements → data assets, involves crucial steps such as data collection, storage, management, analysis, and application. This transforms data into a vital resource for driving innovation and development, enhancing its economic and practical value. The formal implementation in January 2024 of the Interim Provisions on Accounting Treatment for Enterprise Data Resources (hereinafter referred to as the Interim Provisions) signifies the transition of corporate data assets from natural resources to economic assets, fully embodying the process of data resource valorisation. In-depth analysis of the resource and application effects of data assetisation reveals its significant impact on enhancing corporate investment efficiency. On the one hand, from the perspective of resource effects: Data assetisation elevates the economic value of data assets, enabling their entry into data markets as tradable commodities and generating monetary appreciation for enterprises. Concurrently, through further cleansing and enrichment, data assetisation endows data with attributes tailored to specific commercial contexts, enhancing its scarcity. This facilitates credit pledge functions analogous to intangible assets, thereby strengthening the financing capabilities of data assets. The resource effect of data assetisation improves corporate liquidity, alleviating investment shortfalls caused by financing constraints [8]. Fintech plays a pivotal role in this process. By providing advanced technological tools and innovative financial solutions, it accelerates data assetisation, thereby optimising corporate investment efficiency. From an application perspective, data assetisation transforms corporate data resources into valuable assets through processing and value creation, endowing them with resource element attributes. This process continuously incorporates new data resources, enhancing the economic value of data assets while providing management with fresh avenues for obtaining authentic operational insights. Through prudent exploitation of data assets and market forecasting, their supervisory and application functions can be effectively leveraged, thereby mitigating overinvestment stemming from principal-agent issues [9]. FinTech facilitates this process by establishing data asset management systems and providing relevant tools, assisting enterprises in managing the lifecycle of data assets. This enhances operational efficiency, realises value creation, and drives improvements in corporate investment efficiency. In summary, investigating whether data assetisation can enhance corporate investment efficiency, through which pathways it achieves this, and how FinTech development influences these processes holds significant theoretical and practical implications for stimulating stable and orderly economic growth.

 doi:10.20944/preprints202603.0705.v1

This study examines the impact of data assetisation on corporate investment efficiency using Shanghai and Shenzhen A-share listed companies from 2012 to 2023 as the research sample. Firstly, employing the Word2Vec neural network model and deep learning techniques to construct a measurement index for corporate data assetisation, it is found that data assetisation effectively enhances corporate investment efficiency. Secondly, data assets are categorised into self-use and transactional types based on usage patterns, with their respective impacts on investment efficiency examined. Thirdly, the study explores the differential effects of data assets on investment efficiency under varying internal and external conditions, considering firm life cycles, resource types, and bank liquidity heterogeneity. Finally, by identifying specific pathways influencing overinvestment and underinvestment, corresponding recommendations are proposed, revealing that bank fintech exerts a positive moderating effect on corporate investment efficiency. Potential contributions include: (1) Enriching micro-level applications of data assetisation. By examining how data assetisation influences corporate investment behaviour and exploring differential impacts of proprietary versus transactional data assets across distinct investment dilemmas, this research narrows its scope to provide micro-level evidence supporting data assetisation's role in empowering the real economy. (2) Expanding research on mechanisms to enhance corporate investment efficiency. This study identifies pathways through which data assetisation improves investment efficiency by addressing both underinvestment and overinvestment, proposing recommendations to enhance human capital levels, strengthen external oversight, and accelerate digital transformation. (3) Focusing on Fintech Development. By matching corporate data with that of their primary lending banks, this study reveals that advancements in banking fintech exert a positive moderating effect on data assetisation's enhancement of corporate investment efficiency. It further demonstrates differential impacts on alleviating underinvestment and overinvestment, highlighting fintech's significance within the data assetisation process and its complex moderating mechanisms.

2. Materials and Methods

2.1. Theoretical Analysis and Research Hypotheses

The convergence of digital and physical economies drives digital transformation, not only enhancing corporate utilisation of data assets but also providing new avenues for improving capital market transparency. This facilitates greater access to investment information and lower financing thresholds for enterprises through data assetisation at both theoretical and practical levels, thereby significantly influencing corporate investment efficiency. Theoretically, as the digital economy evolves, the asset specificity theory within transaction cost economics suggests that more versatile data assets better serve as collateral, thereby enhancing enterprises' credit access capabilities [9]. Practically, data assetisation captures digital transaction and investment information within capital markets, enabling enterprises to readily identify new investment opportunities. Concurrently, corporate digital transformation leverages the informational effects of data assets to effectively reduce internal and external information asymmetry, strengthen internal information management and risk control capabilities, and mitigate agency problems and managerial overconfidence [10]. The development of the digital economy also presents opportunities for the advancement of fintech within financial institutions such as banks [11]. On the one hand, the application of digital technologies provides convenient tools for financial institutions to better assess the value of data assets, facilitating the resource effects of data assetisation. On the other hand, the interaction between data assetisation and fintech not only enhances the accuracy of market investment forecasting but also, through collaborative supervision between enterprises and banks, leverages the regulatory empowerment of data assets, thereby realising the application effects of data assetisation. Based on this, this paper will focus on the resource effects and application effects of data assetisation to analyse its impact on corporate investment efficiency, as well as the moderating influence of banks' fintech capabilities in this process.

doi:10.20944/preprints202603.0705.v1

2.1.1. Mechanistic Analysis of Data Assetisation on Corporate Investment Efficiency

2.1.1.1. Resource Effects of Data Assetisation

Due to issues such as information asymmetry and capital bias in financing markets, corporate investment behaviour is susceptible to internal selection bias and external financing discrimination, resulting in inefficient investment manifested as underinvestment. Consequently, decision-making biases and liquidity shortages are key factors diminishing investment efficiency [12]. Data assetisation effectively promotes knowledge diffusion through its resource effect, thereby optimising investment project selection and alleviating financing constraints, ultimately enhancing corporate investment efficiency. This is analysed as follows. Firstly, from a knowledge diffusion perspective. Through processes such as data search, cleansing, matching, and processing, data assetisation transforms data elements into highly practical data assets. These assets function as novel internal and external information resources, enabling employees to internalise knowledge and further convert it into employee wisdom [13]. The transmission from data elements to employee wisdom constitutes a non-static, multi-stakeholder transformation process that facilitates the accumulation of employee experience and elevates the enterprise's human capital level. Enhancing the comprehensive capabilities of investment personnel optimises investment selection biases, enabling sound investment decisions. Concurrently, by continually elevating corporate awareness of data assets, these assets can serve as crucial evidence for forecasting investment directions, thereby improving investment rationality, alleviating underinvestment issues, and boosting corporate investment efficiency. On the other hand, from the perspective of alleviating financing constraints. China's financial system lags behind the development of its real economy, rendering corporate financing susceptible to transaction costs [14]. Data assetisation employs digital networks as transactional mediums to realise peer-to-peer data pledge models. Concurrently, enterprises categorise data assets into distinct asset classes, facilitating automated identification during the pledge process. This effectively reduces information search and transaction costs between banks and enterprises, thereby alleviating financing constraints. With the implementation of the Interim Provisions, the process of incorporating data assets into balance sheets is becoming increasingly standardised. This enhances the role of data assets as intangible collateral and, through derivatives such as data trusts and data asset securitisation, enables enterprises to generate additional monetary income. Consequently, it improves corporate liquidity and alleviates the predicament of insufficient investment capital. In summary, this leads to the formulation of the first research hypothesis, H1.

H1. *Data assetisation alleviates corporate investment shortfalls by leveraging resource effects, thereby enhancing corporate investment efficiency.*

2.1.1.2. Application Effects of Data Assetisation

Corporate decentralisation, coupled with challenges in supervision—both difficult and costly—causes investment decisions to diverge from owners' interests, leading to suboptimal investment levels. Furthermore, as investment decision-makers, management's short-termism and irrational behaviour often result in overinvestment. Thus, agency problems and short-term corporate strategies are key drivers of diminished investment efficiency. Data assetisation mitigates agency problems by leveraging application effects and optimises long-term development strategies through driving corporate digital transformation, thereby enhancing investment efficiency. From the perspective of alleviating agency problems:Data assetisation enhances the practical value of data assets by strengthening data element extraction and management capabilities. It fully leverages the role of visualised information transmission, standardises information disclosure methods, and significantly diversifies corporate financing channels. This shifts the initiative in business governance and key innovation back to the enterprise's internal structure [15], thereby reinforcing management's core analytical position. In the process of enhancing operational capabilities and driving corporate growth, management reduces traditional agency problems under decentralised models by increasing personal returns and gaining societal recognition [16]. Concurrently, data assetisation facilitates

 doi:10.20944/preprints202603.0705.v1

shareholder comprehension of corporate production, operations, and investment activities by converting operational information into visualised data assets. This reduces internal oversight costs, aligning agents' personal interests with corporate objectives [17], thereby improving principal-agent relationships and enhancing corporate investment efficiency. From the perspective of driving corporate digital transformation,Data assetisation empowers corporate value creation through application effects, enhancing decision-making quality and efficiency by providing more accurate and comprehensive data. This accelerates the formation of a mutually reinforcing relationship between data assetisation and corporate digital transformation, driving deeper digital development. Simultaneously, digital transformation underscores the significance of data assets. By enhancing information mining efficiency and broadening data sourcing channels, it extends corporate governance boundaries through digital information exchange platforms. This amplifies the supervisory effect of governance entities, exemplified by social media [18], thereby standardising corporate governance practices and elevating long-term development preferences. Furthermore, as a long-term development strategy, digital transformation mitigates short-term investment myopia, aligning investment behaviour with corporate development expectations and improving investment efficiency. In summary, this leads to the formulation of the second research hypothesis, H2.

H2. *Data assetisation alleviates excessive corporate investment by leveraging application effects, thereby enhancing investment efficiency.*

2.1.2. The Moderating Role of Fintech

Under the assumption of perfect competition in neoclassical economic theory, credit is primarily determined by supply and demand. However, China's financing landscape has long been dominated by indirect financing, with banks serving as the principal financial intermediaries. Consequently, information frictions and economic uncertainty influence banks' credit decisions, thereby affecting corporate investment behaviour [19]. Advancements in modern technologies such as artificial intelligence and cloud computing have provided opportunities for fintech-driven financial innovation, serving as a catalyst for enhancing corporate investment efficiency through data assetisation. Firstly, considering the resource effects of fintech empowering data assetisation: fintech enables efficient integration with corporate data assets through technological drivers, accurately estimating their economic value. Enhanced by fintech, banks can now utilise data mining and cleansing to identify not only quantifiable "hard information" such as tangible assets and credit ratings, but also "soft information" like entrepreneurial spirit and commercial creditworthiness. This enables precise assessment of corporate behaviour, effectively alleviating financing constraints [20] and boosting corporate investment efficiency. Secondly, considering the application effect of fintech in empowering data assetisation. Through technologies like blockchain and biometrics, fintech strengthens external oversight bodies' supervision of corporate management behaviour. By horizontally transmitting data assets to other investors, it leverages synergistic regulatory effects to mitigate agency problems. Simultaneously, fintech integrates capital market data to complement corporate data assets, jointly providing precise predictive information for corporate investment. This enhances the rationality of investment decisions and alleviates overinvestment phenomena. Finally, regarding fintech's reduction of banks' risk perception. When banks adopt a cautious lending stance in response to economic uncertainty shocks, this shrinks market lending volumes, increases corporate financing costs, and simultaneously diminishes the pledgeability of corporate data assets, thereby intensifying financing constraints faced by enterprises [21]. The advancement of fintech, through precise forecasting of market fluctuations and the adoption of risk diversification strategies, enhances banks' resilience to risks, reduces their risk perception, increases their willingness to lend, and ensures the smooth progression of corporate investment activities. In summary, this leads to the formulation of the third research hypothesis, H3.

H3. *Fintech exerts a positive moderating effect on the enhancement of corporate investment efficiency through data assetisation.*

2.2. Research Design

2.2.1. Sample Selection and Data Sources

This study employs A-share listed companies from 2012 to 2023 as the research sample, with the following data processing steps: Firstly, observations from the financial and real estate sectors are excluded. Secondly, samples of enterprises in special statuses such as ST, *ST, PT, or suspended listing are removed. Finally, samples with excessive missing values in key variables were excluded. All continuous variables underwent trimming at the upper and lower 1% to mitigate the impact of extreme values on regression results. Firm data originated from the Guotai An database (CSMAR) and Wind database, while word frequency data derived from annual reports of listed companies.

2.2.2. Variable Selection and Model Construction

2.2.2.1. Dependent Variables

Investment Efficiency (Inveff):Following Richardson's methodology [22], this study measures inefficient investment levels using the absolute deviation between a firm's actual investment level and its optimal investment level. The magnitude of this value reflects the severity of inefficient investment, serving as a proxy for investment efficiency. The sign of the residual value indicates overinvestment (over_Inveff) or underinvestment (under_Inveff), with the underinvestment indicator taking the absolute value to standardise the direction of the metric. The greater the values of the overinvestment (over_Inveff) and underinvestment (under_Inveff) indices, the more severe the respective investment imbalances become. The specific model construction is illustrated in Equation (1).

$$\begin{aligned} Inv_{it} = \beta_0 + \beta_1 Grow_{i\,t-1} + \beta_2 Lev_{i\,t-1} + \beta_3 Size_{i\,t-1} + \beta_4 Cash_{i\,t-1} + \beta_5 Age_{i\,t-1} \\ + \beta_6 Ret_{i\,t-1} + \beta_7 Invest_{i\,t-1} + \sum Year + \sum Industry + \varepsilon \end{aligned} \tag{1}$$

where: Inv_{it} denotes the new investment expenditure of firm i in year t, Grow represents the operating revenue growth rate,Lev denotes the firm's leverage ratio, Size represents the firm's asset scale, Cash indicates the firm's cash flow situation, Age signifies the firm's listing duration, Ret denotes the firm's stock return rate,$Invest_{i\,t-1}$ represents the new investment expenditure of firm i in year t-1, Year is the year dummy variable, and Industry is the industry dummy variable. To reflect the lagged impact of firm-specific factors on investment behaviour, all influencing factors are treated with a one-period lag.

2.2.2.2. Explanatory Variables

Data Monetisation (Da): This study references the research by He Ying et al. [5,9], employing text analysis with "information", "network", "digital", and "data" as seed terms. A neural network model and deep learning were used to construct a corpus of similar words. The logarithm of word frequency plus one was adopted as the proxy variable for data assetisation. data assetisation was further categorised into own-use (Oda) and transactional (Bda) types based on data usage.

2.2.2.3. Control Variables

This study selected the following control variables potentially influencing corporate hiring to mitigate the impact of sample selection bias on regression outcomes. These include: Enterprise Ownership Structure (SOE): Set to 1 if the enterprise is state-owned, otherwise 0; Enterprise Scale (Size): Measured using the logarithm of total assets;Debt-to-Asset Ratio (Lev): measured as total liabilities divided by total assets; Total Asset Growth Rate (AssetGrowth): measured as the difference between current year total assets and previous year total assets divided by the previous year's total assets; Board Size (Board): measured as the natural logarithm of the number of board members;Dual Role (Dual): Measured by a dummy variable indicating whether the Chairman and Chief Executive Officer are held by the same individual; set to 1 if so, otherwise 0; Tobin's Q Ratio (TobinQ): Measured by the ratio of the firm's market value to its asset replacement cost;Firm Age: Measured using the logarithm of the difference between the current year and the year of establishment; Employee Count: Measured using the logarithm of the total number of employees; Bank Shareholding: Set to 1 if shares are held in banks or other financial institutions, otherwise 0.

2.2.2.4. Model Construction

To examine the impact of data assetisation on corporate investment efficiency, the following econometric model is established, drawing upon Yu, Minggui et al. [23].

$$Ineff f_{it} = \beta_0 + \beta_1 Da_{i,t} + \beta_z C_{i,t} + r_i + \theta_t + \varepsilon \qquad (2)$$

where $Ineff f_{it}$ denotes the inefficient investment of firm i in year t;$\mathrm{Da}_{i,t}$ represents the data assetisation development of firm i in year t;$\mathrm{C}_{i,t}$ comprises various control variables potentially influencing corporate investment efficiency;r_i andθ_t denote industry fixed effects and time fixed effects respectively, to exclude industry heterogeneity that does not vary over time and reduce interference from time effects;ε constitutes the random disturbance term;

3. Results

3.1. Descriptive Statistics

Table 1 presents descriptive statistics for key variables. The mean investment efficiency (Inveff) is 0.030 with a standard deviation of 0.115, indicating volatility in investment efficiency across sample enterprises. The mean for data assetisation (Da) is 0.734 with a standard deviation of 2.291, indicating significant variation in the degree of data assetisation across firms. This disparity may stem from the differences in resource attributes that influence firms' demand for and application of data assets。. The statistical characteristics of the remaining control variables align with prior research, though notable differences exist between firms. These variables must be incorporated into the regression model to ensure robust results.

Table 1. Descriptive Statistics of Key Variables.

VarName	Obs	Mean	SD	Min	Median	Max
Inveff	29278	0.030	0.115	-1.096	-0.010	1.448
Da	29278	2.291	0.734	0.000	2.197	5.908
SOE	29278	0.355	0.479	0.000	0.000	1.000
Size	29278	22.263	1.296	19.585	22.07	26.452
Lev	29278	0.426	0.204	0.032	0.418	0.908
Asset Growth	29278	0.168	0.368	-0.383	0.088	5.116
Board	29278	2.120	0.197	1.609	2.197	2.708
Dual	29278	0.284	0.451	0.000	0.000	1.000
TobinQ	29278	2.042	1.382	0.802	1.605	15.607
FirmAge	29278	2.927	0.332	1.386	2.995	3.611
Employee	29,278	7,676	1.245	4.174	7.597	11.181
Bank	29,278	0.056	0.230	0.000	0.000	1.000

3.2. Benchmark Regression Analysis

3.2.1. Impact of Data Assetisation on Corporate Debt Concentration and Debt Maturity

The VIF and Hausman tests indicate no significant multicollinearity in the sample data, warranting the use of a fixed-effects model for empirical analysis. To ensure robustness of regression results, randomness was mitigated by progressively incorporating control variables and fixed effects.

 doi:10.20944/preprints202603.0705.v1

The regression results are presented in Table 2. Column (1) shows the results without control variables or fixed effects, where the regression coefficient for Data Assetisation (Da) is significantly negative, preliminarily demonstrating that data assetisation can reduce inefficient investment in enterprises. Column (2) presents the results with fixed effects but without control variables, while Column (3) shows the results with control variables but without fixed effects. In both cases, the regression coefficient for Data Assetisation (Da) is significantly negative. Column (4) presents results with control variables and fixed industry and time characteristics. The regression coefficient for data assetisation (Da) is -0.010, passing the 1% significance test, indicating that data assetisation (Da) effectively enhances corporate investment efficiency. From an economic significance perspective, a one-standard-deviation increase in data assetisation (Da) levels raises corporate investment efficiency by 24.46%, demonstrating substantial economic significance. Hypothesis A1 is validated.

Table 2. Benchmark Regression Results.

VARIABLES	(1) Inveff	(2) Inveff	(3) Inveff	(4) Inveff
Da	-0.008***	-0.011***	-0.006***	-0.010***
	(-8.33)	(-9.17)	(-6.81)	(-8.79)
SOE			0.008***	0.011***
			(5.74)	(6.88)
Size			-0.008***	-0.010***
			(-9.71)	(-10.48)
Lev			0.011***	0.019***
			(3.14)	(5.08)
AssetGrowth			0.143***	0.146***
			(74.60)	(76.11)
Board			-0.000	-0.000
			(-0.14)	(-0.08)
Dual			-0.003**	-0.004***
			(-2.01)	(-2.74)
TobinQ			-0.008***	-0.008***
			(-15.27)	(-14.56)
FirmAge			0.016***	0.007***
			(7.79)	(2.95)
Employee			0.008***	0.008***
			(10.48)	(9.10)
Bank			-0.006**	0.001
			(-2.29)	(0.32)
Constant	0.018***	-0.005	0.071***	0.128***
	(8.16)	(-0.48)	(4.30)	(6.22)
Fixed effects	No	Yes	No	Yes
Observations	29,278	29,278	29,278	29,278
R-squared	0.002	0.020	0.174	0.195

Note: ***, **, * denote significance levels of 1%, 5%, and 10% respectively; values in parentheses indicate t-statistics; same applies below.

3.2.2. Impact of Self-Use Data Assets and Tradeable Data Assets on Corporate Investment Efficiency

To further refine the effects of data assetisation, data assets were categorised by usage into internal-use data assets and transactional data assets. This approach examined the impact of the economic and practical effects of data assetisation on corporate investment efficiency. The regression results are presented in Table 3. The test results in columns (1) and (4) indicate that both internal-use data assets (Oda) and transactional data assets (Bda) significantly reduce corporate inefficient investment and enhance investment efficiency, with internal-use data assets demonstrating a more

 doi:10.20944/preprints202603.0705.v1

pronounced enabling effect. By substituting investment deficiency (under_Inveff) and overinvestment (over_Inveff) for investment efficiency, it is found that proprietary data assets demonstrate a significant advantage in mitigating investment deficiency, while transactional data assets prove more effective in alleviating overinvestment. These findings indicate that proprietary data assets alleviate underinvestment by enhancing corporate utilisation of data resources, improving market forecasting accuracy, and boosting operational efficiency, thereby ensuring sufficient investment capital. Concurrently, transactional data assets mitigate overinvestment by strengthening corporate disclosure practices and reinforcing internal and external collaborative oversight.

Table 3. Regression Results by Data Asset Dimension.

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
VARIABLES	**Inveff**	**over_Inveff**	**under_Inveff**	**Inveff**	**over_Inveff**	**under_Inveff**
Oda	-0.009***	-0.004**	-0.003**			
	(-7.91)	(-2.54)	(-2.37)			
Bda				-0.008***	-0.006***	-0.002*
				(-7.34)	(-3.77)	(-1.95)
Constant	0.126***	0.081***	0.050***	0.091***	0.066***	0.042**
	(6.11)	(3.14)	(2.72)	(4.54)	(2.61)	(2.29)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	29,278	29,278	29,278	29,278	29,278	29,278
R-squared	0.194	0.418	0.103	0.194	0.418	0.102

3.3. Endogeneity Tests

3.3.1. Heckman Two-Step Method

As the indicators for corporate data assetisation are constructed using Word2Vec neural network models and deep machine learning techniques, with corporate annual reports and other publicly available information serving as key references, firms possess an incentive to exaggerate their data assetisation levels to secure greater external funding. This readily introduces bias between the true and estimated values of corporate data assetisation, thereby generating endogeneity issues stemming from sample selection bias. To mitigate the impact of this issue on regression results, this study employs Heckman's two-step method for re-estimation. First, a proxy variable (XR) for corporate data asset disclosure is constructed. This variable is set to 1 if the company's public announcements during the current year contain seed terms such as "information," "network," "digital," or "data," and 0 otherwise. Secondly, using the dummy variable (XR) as the dependent variable and all control variables as covariates, a first-stage regression test is conducted via a Probit model to derive the inverse Mills ratio (IMR). Finally, the IMR is incorporated into the baseline regression model (2) for a second-stage regression test. The results of the second step are presented in Column (1) of Table 4. The coefficient for data assetisation (Da) is -0.010 and passes the 1% significance test, while the IMR coefficient is 0.049 and also passes the 1% significance test. This indicates that the regression results remain robust after overcoming sample selection bias.

3.3.2. Propensity Score Matching (PSM)

The benchmark regression previously incorporated industry and year fixed effects to eliminate the influence of industry characteristics that do not vary over time and temporal characteristics that do not vary across industries. However, firm characteristics exert differential effects on data assetisation development, potentially introducing endogeneity issues due to omitted variables. This paper employs Propensity Score Matching (PSM) to address these issues. First, enterprises are stratified into high- and low-data-assetisation groups by calculating the median data-assetisation score, assigning values of 1 and 0 respectively. Second, all control variables are treated as covariates

 doi:10.20944/preprints202603.0705.v1

for 1:1 propensity score matching. Finally, the matched samples undergo re-testing. The regression results are presented in Column (2) of Table 4. After controlling for omitted variables in the sample, the coefficient for data assetisation (Da) is -0.009 and passes the 1% significance test, consistent with previous research findings.

3.3.3. Dynamic Panel Regression

A firm's investment preferences in the previous period may influence the current management's investment choices, meaning that investment efficiency is affected by prior decisions. This can generate time-series autocorrelation, introducing endogeneity bias. Therefore, this study employs System Generalised Method of Moments (SYS-GMM) to address the endogeneity issue arising from autocorrelation. The regression results are presented in Column (3) of Table 4. The AR(1) coefficient is significant while the AR(2) coefficient is insignificant, indicating that the disturbance term exhibits no autocorrelation issues, thus meeting the requirements for dynamic panel analysis. The data assetisation (Da) coefficient is -0.225 and passes the 1% significance test, demonstrating that after mitigating serial autocorrelation characteristics, data assetisation remains effective in enhancing corporate investment efficiency.

3.3.4. Instrumental Variables Method

The benchmark regression results indicate that data assetisation can effectively enhance corporate investment efficiency. Notably, enterprises with higher investment efficiency possess greater motivation to expand financing scale and further seek investment opportunities in order to fully leverage their investment advantages. This, in turn, imposes higher demands on the quality of collateral and the rationality of investment decisions, potentially further driving the development of corporate data assetisation. Consequently, an endogeneity issue involving reverse causality may exist between data assetisation and corporate investment efficiency. To mitigate biased estimation of regression results due to endogeneity, this study employs two-stage least squares (2SLS) for endogeneity testing. Following the methodology of Xiao, Tu, and Sheng [24] and He and Ying [9], the mean value of data assets for other enterprises in the same industry during the current year is utilised as an instrumental variable (IV). On the one hand, due to cluster effects and competitive pressures, when data assetisation within an industry yields economic benefits for other firms, it intensifies competitive pressures, prompting firms to prioritise data asset development and application—thus satisfying the instrumentality requirement. On the other hand, the data assetisation of other firms within the industry cannot directly influence a firm's own investment efficiency, meeting the exogeneity requirement. The results of the first stage of the 2SLS regression are presented in Table 4. The instrumental variable (IV) regression coefficient is significantly positive. The *Cragg-Donald Wald* F-statistic is 18.35, exceeding the 10% significance threshold of 16.38 for weak instrumental variable tests, thus rejecting the weak IV hypothesis and satisfying the correlation requirement. The *Kleibergen-Paap rk LM* statistic is 12.46 and passes the 1% significance test, rejecting the null hypothesis of unidentifiability of the instrumental variable. The results of the second stage of the 2SLS regression indicate that the regression coefficient for data assetisation (Da) is -0.123 and passes the 5% significance test. This demonstrates that, after controlling for reverse causality interference, data assetisation still enhances corporate investment efficiency. To further validate the exogeneity of the instrumental variables, this study incorporates them into regression model (2), following the methodology of Tang et al. [25]. The regression results are presented in column (6). Here, the regression coefficient for data assetisation (Da) is significantly negative, while the instrument variable (IV) coefficient is insignificant, further confirming that the instrument variable satisfies the exogeneity requirement.

 doi:10.20944/preprints202603.0705.v1

Table 4. Endogeneity Test.

Variable name	(1) Heckman's two-step method Inveff	(2) PSM Inveff	(3) GMM Inveff	(4) Instrumental Variables Da	(5) Instrumental Variables Inveff	(6) Instrumental Variables Inveff
Da	-0.010***	-0.009***	-0.225***		-0.123**	-0.010***
	(-8.41)	(-5.98)	(-20.62)			(-8.75)
IMR	0.049***					
	(4.25)					
IV				0.025***		-0.003
				(4.21)		(-0.07)
AR(1)			-11.54***			
AR(2)			-2.50			
Constant	0.162***	0.081***	-0.251	3.599***	0.458**	0.124***
	(7.44)	(3.05)	(-0.62)	(4.15)	(2.54)	(6.02)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	29,278	16,373	24,782	29,278	29,278	29,278
R-squared	0.190	0.194		0.434	0.590	0.195

3.4. Robustness Tests

3.4.1. Substitution of the Dependent Variable

First, drawing upon Biddle [26]'s methodology, we employ the sales revenue growth rate to gauge corporate income growth levels. The regression residual coefficient of this growth rate serves as a measure of investment efficiency, with higher values indicating lower efficiency. Secondly, drawing upon Chen [27]'s methodology, a NEG dummy variable was incorporated into Biddle's model. This accounts for differing relationships between investment and sales growth rates during periods of rising versus falling revenues. The NEG variable is set to 1 when sales growth declines and 0 otherwise. Investment efficiency is then measured by the absolute value of the model residuals, where a larger residual indicates lower investment efficiency. The regression results are presented in Table 5, Columns (1) and (2). The regression coefficients for data assetisation are significantly negative in both cases, indicating robust results.

3.4.2. Addressing Serial Correlation Issues

First, to mitigate the impact of heteroskedasticity and serial correlation on regression results, this study employs Driscoll and Kraay robust standard errors for re-testing. By individually standardising sample units and recombining them into new panel data, individual variations are eliminated to enhance estimation accuracy. Test results (Column 3, Table 5) show significantly negative data assetisation regression coefficients, confirming robustness.

Secondly, to further address serial correlation, this study incorporates data assetisation from both preceding and subsequent periods into the regression. This mitigates the impact of temporal dependencies on regression outcomes. Results indicate that regression coefficients for both ex ante (Da) and ex post data assetisation prove insignificant, with only the current period yielding a significant coefficient. This confirms that regression results remain robust after controlling for serial correlation.

3.4.3. Sub-Sample Regression

Firstly, enterprises located in economically developed regions possess stronger foundations for data assetisation and more favourable data resource endowments, which may lead to "spurious regression" caused by outliers. Therefore, this paper excludes samples from municipalities directly under the central government and provincial capitals to mitigate the influence of regional comprehensive endowments on regression results. Secondly, in September 2015, the State Council issued the Action Plan for Promoting Big Data Development, signifying big data's elevation to a national-level development strategy and substantially enhancing the importance of data as an economic factor. Consequently, this study employs sample data from 2015 to 2023 to minimise the impact of missing data assetisation values on regression outcomes. The test results, as shown in columns (5) and (6) of Table 5, indicate that the regression coefficients for data assetisation are all significantly negative, demonstrating the robustness of the findings.

Table 5. Robustness Test.

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	Replace the dependent variable		**Robust standard error correction**	**Exclusion of serial correlation**	**Subsample regression**	
Variable names	**Inveff**	**Inveff**	**Inveff**	**Inveff**	**Inveff**	**Inveff**
Da	-0.002***	-0.001**	-0.010***	-0.004**	-0.010***	-0.011***
	(-3.06)	(-2.55)	(-3.73)	(-2.07)	(-6.44)	(-9.02)
Constant	0.099***	0.093***	0.128**	0.133***	0.164***	0.102***
	(10.42)	(9.96)	(2.96)	(5.75)	(5.65)	(4.33)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	29,278	29,278	29,278	22,401	15,265	20,973
R-squared	0.235	0.227	3,960	0.197	0.200	0.211

3.5. Heterogeneity Analysis

3.5.1. Corporate Life Cycle

The life cycle theory posits that differing developmental stages of enterprises influence their endowment of data elements, risk preferences, and investment directions. These stage-specific characteristics impact the efficacy of data assetisation on corporate investment efficiency. Drawing upon Dickinson [28], this study employs the cash flow approach to categorise enterprises into three groups: growth, maturity, and decline. Using panel regression, we examine how business life cycle heterogeneity affects the impact of data assetisation on enhancing investment efficiency. The regression results, presented in Table 6 (Columns 1–3), show that the regression coefficients for data assetisation (Da) are all significantly negative. This indicates that data assetisation enhances investment efficiency for enterprises across all life cycle stages. Notably, the between-group difference coefficient is significant, and the regression coefficients follow the order: growth stage < mature stage < decline stage. This suggests that data assetisation exerts a more pronounced promotional effect on investment efficiency for enterprises in the growth stage.

To further explore how corporate life-cycle heterogeneity influences data assetisation's enhancement of investment efficiency, investment efficiency was disaggregated into underinvestment and overinvestment. Test results are presented in columns (4) to (9) of Table 6. Specifically: - Columns (4) to (6) examine the effect of data assetisation (Da) on overinvestment. The coefficient for growth-stage enterprises is significantly negative, indicating that data assetisation most effectively mitigates overinvestment in this stage. Columns (7) to (9) present the test results for data assetisation (Da) on underinvestment. The coefficient for data assetisation in growth-stage enterprises is significantly negative, indicating that data assetisation is more effective in mitigating underinvestment within this phase. This stems from the fact that for enterprises in the growth phase, data assetisation better communicates their digital maturity externally, thereby securing critical resources. Concurrently, it effectively drives a shift in management decision-making from experience-driven to data-driven approaches, thus resolving underinvestment issues. Furthermore, data assetisation enables comprehensive integration of data resources, refining classifications of existing data assets and enhancing inter-firm data exchange. By increasing information transparency,

 doi:10.20944/preprints202603.0705.v1

it mitigates information asymmetry, strengthens oversight of management investment behaviour, and effectively curbs excessive investment.

3.5.2. Types of Enterprise Resources

Based on the intensity of production factors, enterprises can be categorised as labour-intensive, technology-intensive, or asset-intensive. Different enterprise types exhibit significant variations in their emphasis on and application of data assetisation. This paper references the National Economic Industry Classification, designating agriculture, forestry, animal husbandry, and fisheries; manufacturing; commerce, catering, and other services as labour-intensive enterprises. High-tech services such as information services, R&D, and design services are classified as technology-intensive enterprises, while heavy industry and construction are categorised as asset-intensive enterprises. A grouped regression approach was employed to examine how enterprise resource type heterogeneity influences the impact of data assetisation on investment efficiency. The test results, as shown in Table 7, columns (1) to (3), reveal that the regression coefficients for data assetisation (Da) are all significantly negative, with the order being technology-intensive < asset-intensive < labour-intensive. This indicates that data assetisation is most effective in enhancing investment efficiency for technology-intensive enterprises. This may stem from the fact that technology-intensive enterprises typically operate within high-tech service sectors, where data resources form the foundation for enhancing service quality and fostering innovation. Their emphasis on data collection and data assetisation fully unlocks the potential of data dividends to empower investment efficiency gains. Asset-intensive enterprises typically rely on substantial asset investments to sustain daily operations, with assets primarily comprising fixed assets, intangible assets, accounts receivable, and so forth. Data assetisation optimises the investment environment by stabilising operational conditions through enhanced asset management and improving resource utilisation efficiency. Conversely, labour-intensive enterprises exhibit lower digitalisation levels and have yet to leverage the informational advantages of data assets in investment decision-making.

3.5.3. Bank Liquidity

From a banking perspective, when systemic liquidity is ample, banks' capital surpluses significantly increase, effectively meeting lending and investment demands. At such times, the banking system may be more inclined to accept data assets as collateral, alleviating corporate investment shortfalls. From the data assetisation perspective, this transformation constitutes a critical step in enhancing the value of corporate data assets. This process requires financial institutions to confer financial attributes upon such assets, meeting the criteria for high-quality collateral and alleviating the predicament of diminished investment efficiency caused by financing constraints. Consequently, variations in bank liquidity exert heterogeneous effects on data assetisation's capacity to enhance corporate investment efficiency. Following the methodology of Huang Zhen et al. [29], this study employs the year-on-year growth rate of M2 to measure banking system liquidity, where a higher growth rate indicates a more accommodative monetary policy environment and stronger bank liquidity. By selecting the median value of M2 growth rate to divide the observation sample into high-liquidity and low-liquidity groups, a grouped regression approach is adopted to verify the heterogeneous impact of bank liquidity. Test results are presented in Table 7, columns (4) and (5). The regression coefficients for data assetisation (Da) are consistently negative and higher in high-liquidity than low-liquidity regions, indicating that data assetisation more effectively optimises corporate investment in areas with greater banking liquidity. This aligns with the aforementioned hypothesis.

Table 6. Analysis of Lifecycle Heterogeneity.

	(1) Growth stage	(2) Maturity phase	(3) Decline phase	(4) Growth phase	(5) Maturity Stage	(6) Decline phase	(7) Growth phase	(8) Maturity phase	(9) Decline phase
Variable Name	Inveff	Inveff	Inveff	over_Inveff under_Inveff	over_Inveff under_Inveff	over_Inveff under_Inveff	under_Inveff	under_Inveff	under_Inveff
Da	-0.021***	-0.011***	-0.005***	-0.021***	-0.004	-0.001	-0.004**	-0.002	-0.001
	(-5.07)	(-5.79)	(-3.96)	(-4.04)	(-1.56)	(-0.29)	(-2.39)	(-0.62)	(-1.10)
Constant	0.284***	0.154***	0.038	0.161	0.079*	-0.007	0.239***	0.123***	0.027
	(3.67)	(4.19)	(1.56)	(1.54)	(1.81)	(-0.22)	(3.30)	(3.59)	(1.24)
Intergroup difference coefficient		-0.007***			/			/	
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5340	9,669	14,269	5340	9,669	14,269	5,340	9,669	14,269
R-squared	0.371	0.118	0.120	0.517	0.328	0.278	0.173	0.103	0.137

Table 7. Analysis of Corporate Resource Types and Bank Liquidity Heterogeneity.

	(1) Labour-intensive	(2) Technology -intensive	(3) Asset-intensive	(4) High bank liquidity	(5) Low bank liquidity
Variable name	Inveff	Inveff	Inveff	Inveff	Inveff
Da	-0.007***	-0.012***	-0.008**	-0.011***	-0.009***
	(-3.44)	(-8.24)	(-2.50)	(-5.56)	(-6.66)
Constant	0.123***	0.257***	-0.060	0.099***	0.169***
	(3.73)	(9.63)	(-1.21)	(3.89)	(4.74)
Intergroup difference coefficient		-0.006**		-0.001*	
Control Variables	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	10,238	12,565	6,475	18,850	10,428
R-squared	0.193	0.197	0.240	0.195	0.215

4. Discussion

4.1. Mechanism Testing

This paper examines the pathways through which data assetisation influences corporate investment efficiency by testing its mechanism regarding overinvestment and underinvestment. To verify the relevant mechanisms, an econometric model is constructed based on Jiang Ting [30], as follows:

$$M_{it} = \gamma_0 + \gamma_1 Da_{i,t} + \gamma_z C_{i,t} + r_i + \theta_t + \varepsilon_{i,t} \quad (3)$$

Equation (3) represents the linear regression equation of data assetisation (Da) on the mechanism variable M. When validating the underinvestment mechanism, M denotes the firm's financing constraints and human capital level. When validating the overinvestment mechanism, M denotes the firm's agency costs and digital development level. The remaining letter variables retain the same meanings as in Equation (3) above.

4.1.1. Testing the Underinvestment Mechanism

Scholars have suggested that low corporate investment efficiency stems from underinvestment and overinvestment, the former primarily driven by financing constraints and investment project selection bias [7]. On one hand, credit discrimination by financial institutions exacerbates the misallocation of credit resources, leaving most enterprises with insufficient investment liquidity. On the other hand, investment decision biases arising from inadequate professional expertise among investment personnel further intensify the shortage of effective investment, constraining corporate development. Data assetisation alleviates corporate underinvestment and enhances investment efficiency by strengthening information disclosure and leveraging talent attraction. Specifically: first,

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2026 doi:10.20944/preprints202603.0705.v1

we examine the impact of data assetisation on underinvestment, where higher underinvestment index values indicate more severe underinvestment. Second, we employ the SA index and FC index to represent firms' financing constraint levels, with higher values indicating greater constraints. Additionally, we measure corporate investment liquidity using the ratio of current assets to total assets. Finally, drawing upon Huang Zhuo et al. [31], human capital levels are measured by the proportion of employees holding bachelor's degrees or higher, with further analysis of managerial educational backgrounds. The test results are presented in Table 8. Column (1) shows that the regression coefficient for data assetisation (Da) is significantly negative, indicating that data assetisation can effectively alleviate the problem of insufficient investment. Columns (2) to (4) demonstrate that data assetisation can effectively alleviate corporate financing constraints and further increase the ratio of current assets to total assets. By enhancing short-term liquidity, it alleviates the problem of insufficient investment. Columns (5) and (6) reveal a significantly positive coefficient for data assetisation (Da), indicating its role in attracting talent and enhancing investment efficiency. The findings suggest data assetisation more readily elevates human capital quality, thereby alleviating corporate investment shortfalls.

To further examine the enabling effect of banking fintech development on data assetisation's mitigation of investment shortfalls, this study re-examines the aforementioned path mechanisms using the interaction term between data assetisation (Da) and fintech (Fin). The regression results are presented in Table 9. Within the financing constraint pathway, the interaction coefficient is significantly positive and exhibits a higher absolute value than the corresponding coefficient in Table 8. This indicates that the development of banking fintech effectively amplifies the mitigating effect of data assetisation on financing constraints, further enhancing corporate capital adequacy and current asset ratios, thereby facilitating improved investment efficiency. However, regarding the human capital upgrading pathway, fintech development does not demonstrate a further enhancing effect on the quality of employees and management.

Table 8. Analysis of the Underinvestment Mechanism.

Variable Name	(1)	(2)	(3)	(4)	(5)	(6)
		Financing constraints			Human Capital	
	Underinvestment	SA Index	FC Index	Current Ratio	Level of Human Capital	Management Qualifications
Da	-0.002**	-0.002**	-0.004***	0.013***	0.702***	0.040***
	(-2.52)	(-2.27)	(-3.48)	(8.15)	(11.41)	(4.15)
Constant	0.051***	-2.552***	4.410***	0.912***	-41.180***	-2.900***
	(2.73)	(-128.57)	(192.60)	(30.78)	(-36.16)	(-15.99)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	29,278	29,278	29,278	29,278	29,278	29,278
R-squared	0.103	0.821	0.815	0.378	0.330	0.068

Table 9. Testing the Enabling Effects of Fintech.

Variable Name	(1)	(2)	(3)	(4)	(5)
	Financing constraints			Human capital	
	SA Index	FC Index	Current Assets Ratio	Level of Human Capital	Management Qualifications
Da×Fin	-0.003**	-0.024***	0.024***	0.083***	0.015***
	(-2.18)	(-5.89)	(10.69)	(2.77)	(2.74)
Constant	-2.540***	3.768***	0.865***	-41.241***	-1.631***
	(-80.17)	(9.09)	(16.54)	(-23.11)	(-5.04)
Control variables	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes

 doi:10.20944/preprints202603.0705.v1

Observations	29,278	29,278	29,278	29,278	29,278
R-squared	0.850	0.595	0.382	0.341	0.088

4.1.2. Testing the Overinvestment Mechanism

Corporate overinvestment primarily stems from principal-agent problems and managerial irrationality. On one hand, the separation of ownership and management rights shifts managers' focus from maximising shareholder value to maximising personal gain, leading them to excessively pursue high-profit, high-risk investment projects and exacerbating overinvestment [20]. On the other hand, excessive optimism among managers regarding macroeconomic and industry prospects may induce herd behaviour in investment decisions, lacking scientific rigour. Concurrently, overconfidence within management leads to neglect of investment signals from peers, resulting in overlapping investment domains that exacerbate overinvestment and ultimately diminish corporate investment efficiency [6]. Data assetisation mitigates corporate overinvestment and enhances investment efficiency by strengthening internal and external oversight while accelerating digital transformation. Specifically: firstly, following the methodology of ANG J S [32], agency costs are measured by the ratio of total sales and administrative expenses to operating revenue, with higher values indicating more severe principal-agent issues. Secondly, drawing upon the research of Wu Fei et al. [33], corporate digitalisation levels are assessed using the logarithm of word frequency,where a higher value indicates greater digitalisation. Additionally, following Jiang, Fu Xiu et al. [34], we measure managerial overconfidence through relative executive compensation. If the combined remuneration of the top three executives exceeds the sample mean of executive pay, the firm is deemed to exhibit managerial overconfidence and assigned a value of 1; otherwise, it is assigned 0. The regression results are presented in Table 10. Column (1) shows a significantly negative regression coefficient for data assetisation (Da), indicating that data assetisation effectively mitigates overinvestment issues. Columns (2), (4), and (6) demonstrate that data assetisation drives digital transformation by leveraging synergistic regulatory effects to alleviate agency problems and managerial overconfidence, thereby reducing corporate overinvestment and enhancing investment efficiency. In summary, Hypothesis A2 is validated.

To further examine the enabling role of banks' fintech development in mitigating overinvestment through data assetisation, the interaction term between data assetisation (Da) and fintech (Fin) was substituted for the original explanatory variable in the model. The test results are presented in Columns (3), (5), and (7) of Table 10. Specifically, while fintech development effectively enhances data assetisation's mitigation of agency problems, it exerts a crowding-out effect on data assetisation's capacity to elevate corporate digitalisation levels and curb managerial overconfidence. This may stem from fintech's dual impact: while facilitating convenient financing channels, it simultaneously intensifies the trend of enterprises "shifting away from the real economy towards the virtual", thereby suppressing digital transformation efforts. In mitigating managerial overconfidence through data assetisation, fintech may paradoxically induce renewed "data illusions". Blind optimism regarding data volume and data-driven biases can exacerbate excessive investment in certain areas, thereby diminishing corporate investment efficiency.

Table 10. Analysis of Overinvestment Mechanisms.

Variable Name	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Agency Problem		Digital Transformation			
	Overinvestment	Agency Costs	Agency Costs	Digitalisation Level	Digitalisation Level	Overconfidence	Overconfidence
Da	-0.005***	-0.007***		0.741***		-0.388**	
	(-3.30)	(-4.31)		(4.37)		(-2.41)	
Da×Fin			-0.010**		-0.120**		0.188***
			(-6.52)		(-2.18)		(3.71)
Constant	0.085***	0.320***	0.278***	122.923***	118.558***	108.628***	108.581***
	(3.29)	(11.53)	(10.96)	(39.41)	(20.98)	(36.51)	(20.68)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	29,278	29,278	29,278	29,278	29,278	29,278	29,278
R-squared	0.418	0.082	0.218	0.979	0.961	0.348	0.332

4.1.3. The Moderating Effect of Banking Fintech

To examine the moderating role of banking fintech levels in enhancing corporate investment efficiency through data monetisation, this study establishes the testing model as shown in Equation (4).

$$Inveff_{it} = \beta_0 + \beta_1 Da_{i,t} + \beta_2 Da_{i,t} \times Fin_{i,t} + \beta_3 Fin_{i,t} + \beta_z C_{i,t} + r_i + \theta_t + \varepsilon \quad (4)$$

where $Fin_{i,t}$ denotes the fintech development status of banks providing credit services to firm i in year t. Drawing upon the methodology of Jin Hongfei et al. [35], this study constructs a fintech lexicon for banks across four dimensions: digitalisation, informatisation, internetisation, and intelligentisation. Python is employed to crawl corresponding term frequency data from Baidu search results and bank annual reports, with the level of bank fintech measured by the logarithm of the sum of term frequencies across these four dimensions. The key coefficient examined isβ_2. A significantly negativeβ_2 indicates that the bank's fintech development positively moderates the enhancement of corporate investment efficiency through data assetisation. The regression results are presented in Table 11. Column (1) shows a cross-term coefficient of -0.007, passing the 1% significance test, indicating that the bank's fintech development positively promotes investment efficiency through corporate data assetisation. This may stem from fintech's utilisation of emerging digital technologies such as big data and cloud computing to achieve efficient collection and analysis of market and corporate data. This enables effective assessment of the economic value of corporate data assets and ensures investors can promptly access critical analytical data. Consequently, it alleviates financing constraints while enhancing the scientific rigour of corporate investment decisions. To further validate the direction of fintech's influence, this study replaces corporate investment efficiency with underinvestment and overinvestment for testing. The regression results are shown in Column (2) and Column (3). Column (2) exhibits a significantly negative interaction coefficient, while Column (3) shows a significantly positive interaction coefficient. This indicates that fintech positively promotes the mitigation of underinvestment through corporate data assetisation, yet exerts a suppressing effect on alleviating corporate overinvestment. This may stem from fintech enhancing market transparency to provide diversified investment support for enterprises, thereby economically empowering data assets and better alleviating funding shortages. However, fintech development also heightens the risk of enterprises "shifting away from the real economy towards the virtual economy," exacerbating overinvestment.

Table 11. Moderating Effect of Fintech.

Variable Name	**(1)**	**(2)**	**(3)**
		Bank Fintech Level	
	Inveff under_Inveff	**under_Inveff**	**over_Inveff under_Inveff**
Da	-0.014*	-0.020**	-0.004*
	(-1.92)	(-2.15)	(-1.70)
Da×Fin	-0.007***	-0.008***	0.007***
	(-4.22)	(-5.65)	(3.19)
Fin	-0.003	-0.005	-0.001
	(-0.73)	(-1.06)	(-0.36)
Constant	0.088**	0.040	0.057*
	(2.29)	(0.82)	(1.65)
Control variables	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes
Observations	29,278	29,278	29,278
R-squared	0.197	0.404	0.114

 doi:10.20944/preprints202603.0705.v1

4.2. Conclusions from Mechanism Testing

The results of mechanism testing indicate that data assetisation can effectively enhance corporate investment efficiency through multiple pathways. Specifically: Firstly, regarding the mechanism for alleviating underinvestment, data assetisation significantly reduces financing constraints by leveraging the 'resource effect', thereby increasing the current ratio and enhancing liquidity of corporate funds. Concurrently, through the 'talent effect', it elevates overall human capital levels and improves the educational structure of management teams, optimising the rationality of investment decisions. This effectively mitigates underinvestment stemming from capital shortages or decision-making biases. Second, regarding the mechanism for curbing excessive investment, data assetisation significantly reduces agency costs through its 'application effect', curbing management overconfidence. By driving digital transformation, it enhances information transparency and decision-making rationality, thereby effectively restraining excessive investment stemming from agency problems or irrational behaviour. In summary, data assetisation possesses dual functions of 'resource empowerment' and 'governance empowerment'. It synergistically addresses both underinvestment and overinvestment, ultimately achieving an overall enhancement in corporate investment efficiency.

5. Conclusions

This study examines the impact of data assetisation on corporate investment efficiency using a sample of A-share listed companies on the Shanghai and Shenzhen stock exchanges from 2012 to 2023. Findings indicate that: (1) Data assetisation significantly enhances corporate investment efficiency. This conclusion remains valid after endogenous and robustness tests. Further categorisation of data assets into self-use and transactional types reveals that self-use data assets yield more pronounced effects. (2) Heterogeneity tests reveal that data assetisation exerts a stronger positive effect on investment efficiency for firms in the growth stage, technology-intensive enterprises, and companies operating in regions with high bank liquidity. (3) Mechanism tests indicate that data assetisation enhances investment efficiency by alleviating financing constraints and elevating human capital levels to reduce underinvestment. Concurrently, it mitigates excessive investment by easing agency problems and accelerating digital transformation. (4) Matching firm data with primary lending banks reveals that banks' fintech capabilities positively moderate the effect of data assetisation on investment efficiency. This moderation enhances the mitigation of underinvestment but dampens the reduction of overinvestment. In summary, the following recommendations are proposed.

First, for governmental and supervisory bodies, national-level strategic planning should be implemented. Continuous oversight is required to ensure the standardised accounting treatment of corporate data resources is properly implemented. Guidance should be provided to enterprises on accurately disclosing data asset information, thereby enhancing the transparency and credibility of corporate data assets. This will bolster investor confidence and promote improvements in corporate investment efficiency. Concurrently, the government should strengthen support for SMEs implementing digital strategies, injecting vitality into investment. For enterprises financing through data assets, governments should offer the tax and fee reduction policies or other incentive measures to support investment in data assetisation, lower financing thresholds, and alleviate underinvestment.

Secondly, financial institutions such as banks should proactively respond to corporate financing demands by developing targeted innovations in data asset financial services systems to better assess the value of corporate data assets. Banks and other financial institutions should establish data asset catalogues, prioritise the development of fintech, and enhance the efficiency of interfacing with corporate data assets. This will alleviate corporate funding shortages and improve the scientific rigour of investment decision-making. Concurrently, they should seize opportunities presented by the digital economy to drive their own digital transformation, pioneering new data asset valuation

 doi:10.20944/preprints202603.0705.v1

and risk assessment services. This approach will expand their own economic growth points while providing convenient financial support for corporate investment.

Thirdly, enterprises should actively advance the development of data assets, particularly the cultivation and utilisation of self-consumed data assets, thereby improving investment efficiency. Concurrently, enterprises should prioritise digital development by establishing and refining data asset management platforms to supplement fundamental data asset elements. Fully leveraging the attraction of digital transformation to human capital, enhancing employees' digital literacy and management's professional investment rationality will lay the "human" and "financial" foundations for anchoring emerging investment directions.

References

1. Ye, Y.W.; Yu, T.T.; Tao, Y.Q.; et al. The Investment Stabilisation Effect of Government Data Element Sharing: Evidence from Public Data Open Platforms. Journal of Quantitative Economics and Technical Economics 2024, 1-21.
2. Chen, D.Q.; Li, S.F.; Lei, G.Y. Government Governance, Control Structure and Investment Decisions: Empirical Evidence from Family-Owned Listed Companies. Journal of Finance and Banking 2012, (03), 124-138.
3. Chen, D.Q.; Li, S.F.; Zhong, Y.J. Government Quality, Investment and Capital Allocation Efficiency. World Economy 2012, 35(03), 89-110.
4. Zhao, C.; Su, L.S.; Cao, W. Mixed Ownership Reform: Governance Effect or Resource Effect? — A Study on Corporate Investment Efficiency under Different Property Rights Regimes. Journal of Shanghai University of Finance and Economics 2021, 23(01), 75-90.
5. Yi, C.J.; Zhao, X.Y. Can Digital Transformation Enhance the Overseas Investment Efficiency of Chinese Multinational Enterprises? China Industrial Economics 2024, (01), 150-169.
6. Xia, H.C.; Yang, C. Managerial Overconfidence, Management Autonomy and Corporate Investment Efficiency. Studies in Science of Science 2024, 1-20.
7. Pan, Y.; Tang, X.D.; Ning, B.; et al. Chain Shareholders and Corporate Investment Efficiency: Governance Synergy or Competitive Collusion. China Industrial Economics 2020, (02), 136-164.
8. Wang, Y.; Yang, D. Transformation of the Chinese Management Accounting System: From Data Elements to Data Assets. Management World 2024, 40(10), 171-189.
9. He, Y.; Chen, L.L.; Du, Y.G. Can Data Assetisation Alleviate Financing Constraints for Specialised, Sophisticated, Distinctive and Innovative SMEs? China Industrial Economics 2024, (08), 154-173.
10. Sun, H.B.; Cao, D.; Liu, Z.L. Climate Policy Uncertainty, Digital Transformation and Corporate Investment Efficiency. Journal of Shanghai University of Finance and Economics 2024, 26(06), 62-77.
11. Gu, H.F.; Shen, Y. Does the Development of Digital Finance Enhance Corporate Investment Efficiency? Evidence from Chinese A-share Listed Companies. Journal of Hunan University (Social Sciences Edition) 2024, 38(06), 51-62.
12. Zhao, Z.Y.; Wang, Y.T.; Chen, J.N. Non-Administrative Penalty Regulation and Corporate Investment Efficiency: Empirical Evidence from Exchange Inquiry Letters. Nankai Economic Research 2022, (05), 181-200.
13. Yuan, Z.M.; Yu, X.; Li, M.; et al. Does Data Assets Enhance Human Capital Levels in Chinese Enterprises? Empirical Evidence Based on Text Analysis. Nankai Management Review 2024, 1-27.
14. Zhang, J.; Lu, Z.; Zheng, W.P.; et al. Financing Constraints, Financing Channels and Corporate R&D Investment. World Economy 2012, 35(10), 66-90.
15. Zheng, Z.G.; Li, M.; Yong, H.Y.; et al. Does Shareholder Concerted Action Improve Corporate Governance? Research in Finance 2022, (05), 152-169.
16. Zhu, Q.Y.; Tan, X.F. Corporate Digital Transformation and Investment Efficiency: Empirical Evidence from Annual Report Texts. Economic Science 2024, (04), 93-113.
17. Liu, Z.W.; Zhang, Q.S. Research on Data Element Empowerment of Corporate Investment Efficiency: A Quasi-Natural Experiment Based on Public Data Openness. Economist 2024, (09), 42-52.
18. Yermack, D. Corporate governance and blockchains. Review of Finance 2017, 21(1), 7-31.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2026 doi:10.20944/preprints202603.0705.v1

19. Feng, Y.Q.; Cao, Y. Bank Fintech, Economic Policy Uncertainty and Corporate Investment Behaviour. Contemporary Finance and Economics 2024, (05), 58-72.
20. Li, Z.H.; Zhang, X.D. Can Banking Fintech Improve Corporate Investment Efficiency? Evidence from Chinese Listed Companies. Journal of Nankai University (Philosophy and Social Sciences Edition) 2024, (05), 101-113.
21. Feng, Y.Q.; Zhang, H.L.; Ni, J. Financial Asset Investment Motives of Real Economy Enterprises: The Heterogeneous Impact of Monetary Policy and the Moderation Effect of Digital Finance. China Industrial Economics 2024, (02), 118-136.
22. Richardson, S. Over-investment of free cash flow. Review of Accounting Studies 2006, 11, 159-189.
23. Yu, M.G.; Wang, K. Local Government Debt Financing, Crowding-Out Effects and Corporate Labour Employment. Economic Research Journal 2022, 57(02), 58-72.
24. Xiao, T.S.; Sun, R.Q.; Yuan, C.; et al. Corporate digital transformation, human capital restructuring and the labour income share. Management World 2022, 38(12), 220-237.
25. Tang, Y.J.; Wang, Y.; Tang, C.H. Digital Economy, Market Structure and Innovation Performance. China Industrial Economics 2022, (10), 62-80.
26. Biddle, G.C.; Hilary, G.; Verdi, R.S. How does financial reporting quality relate to investment efficiency? Journal of Accounting and Economics 2009, 48(2-3), 112-131.
27. Chen, F.; Hope, O.K.; Li, Q.; et al. Financial reporting quality and investment efficiency of private firms in emerging markets. The Accounting Review 2011, 86(4), 1255-1288.
28. Dickinson, V. Cash flow patterns as a proxy for firm life cycle. The Accounting Review 2011, 86(6), 1969-1994.
29. Huang, Z.; Zhou, X.J.; Guo, Y. The Debt Structure Channel in the Central Bank's Collateral Management Framework: Empirical Evidence from Bond-Issuing Enterprises. Finance and Trade Economics 2024, 1-17.
30. Jiang, T. Mediating and Moderating Effects in Empirical Studies of Causal Inference. China Industrial Economics 2022, (05), 100-120.
31. Huang, Z.; Tao, Y.Q.; Liu, Z.D.; et al. Smart Manufacturing, Human Capital Upgrading and the Labour Income Share of Enterprises. Economics (Quarterly) 2024, 24(05), 1412-1427.
32. Ang, J.S.; Cole, R.A.; Lin, J.W. Agency costs and ownership structure. The Journal of Finance 2000, 55(1), 81-106.
33. Wu, F.; Hu, H.Z.; Lin, H.Y.; et al. Corporate Digital Transformation and Capital Market Performance: Empirical Evidence from Stock Liquidity. Management World 2021, 37(07), 130-144+10.
34. Jiang, F.X.; Zhang, M.; Lu, Z.F.; et al. Managerial Overconfidence, Corporate Expansion and Financial Distress. Economic Research Journal 2009, 44(01), 131-143.
35. Jin, H.F.; Li, H.J.; Liu, Y.L. FinTech, Banking Risk and Market Crowding-Out Effects. Journal of Finance and Economics 2020, 46(05), 52-65.