Proceedings of the 8th International Conference on Economic Management and Green Development
DOI: 10.54254/2754-1169/104/20242686

A Study of the Impact of Data Assets on the Investment Efficiency of Firms

Yangyi Zhao[1,a], Zixuan Tu[2,b,*], Siyi Chen[3,c]

[1]*Sichuan Normal University, No.1819, Sec.2, Chenglong Avenue, Longquanyi District, Chengdu, Sichuan, China*
[2]*Chengdu University of Technology, No.1, East Third Road, Erxianqiao, Chengdu, Sichuan, China*
[3]*Southwest Petroleum University, No.1, Section 2, Youyuan Road, Shunqing District, Nanchong City, Sichuan Province, China*
a. 17208263019@163.com, b. 1228481753@qq.com, c. 2920809140@qq.com
**corresponding author*

Abstract: With the advent of the digital economy era, data assets play a pivotal role for enterprises. Therefore, this paper selects the data of Chinese A-share listed companies from 2013-2022 to conduct a quantitative study on data assets and explore the impact of enterprises' data assets on their investment efficiency. It is found that data assets significantly contribute to firms' inefficient investment. This finding still holds after the robustness test. Heterogeneity analysis reveals that data assets can promote overinvestment in non-state-owned firms; they contribute more to inefficient investment in growing and declining firms. The mechanism test suggests that data assets promote firms 'inefficient investment due to promoting firms' financing constraints and real earnings management. Therefore, the findings of this paper provide certain clues and suggestions for enterprises to dialectically view data assets to maximize their effectiveness and avoid data bubbles.

Keywords: data assets, investment efficiency, financing constraints, real earnings management.

1. Introduction

In recent years, China's digital economy, which covers a wide range of fields such as cloud computing, big data, and artificial intelligence, has been booming. Currently, there is a large body of literature on the recognition, measurement, and valuation of data assets, or exploring the impact of data assets on the quality of enterprise development or innovation ability. The purpose of this paper is to explore the role of data assets on the investment efficiency of enterprises through empirical analyses.

1.1. Current Status of Research Related to Data Assets

Data assets as an important asset in the era of the digital economy. Lu Zheng et al.[1] argued that data assets can alleviate the problems of information asymmetry and high costs faced by enterprises in the process of development and improve their innovation ability. Li Jian et al. [2]found that data assets can increase the willingness of enterprises to innovate by alleviating their financing constraints, and then increase their investment in innovation. Lu Minfeng[3] pointed out that the data assets of

DOI: 10.54254/2754-1169/104/20242686

enterprises can improve the quality of enterprise development by improving the efficiency of resource allocation. However, some studies also found that data assets are replicable, Jones and Tonett[4] argued that some of the data are non-competitive and can be widely used, so it may undermine the creativity of the firms. Hajli et al.[5]also confirmed that the digitalization of firms can only enhance the performance of industries that are highly dependent on information technology rather than traditional industries.

1.2. Influencing Factors of Enterprise Investment Efficiency

Investment decision-making is widely recognized as an important aspect closely related to the overall strategy and long-term development of an enterprise. Efficient investment is a key driver in determining the high-quality development of an enterprise[6]. However, financing constraints, agency problems, and information asymmetry can lead to lower investment efficiency of enterprises. Therefore, this paper takes this as an entry point to explore the impact of data assets on firms' investment efficiency.

Existing studies have found that with the advancement of digitization, the agency cost problem of managers can not only be alleviated[7], but the integration of digitized information into the development of the enterprise can also reduce the pressure of financing and enhance the investment efficiency of the enterprise[8]. However, some studies have also shown that both accounting standards and policies related to data assets are inadequate[9], which can easily lead to information asymmetry; or make enterprises more complex[10], which leads to earnings management problems. However, there are relatively few studies on the impact and relationship of data assets on corporate investment efficiency, and this paper aims to fill this research gap through research.

2. Rationale and Research Hypothesis

2.1. Data Assets Reduce the Efficiency of Business Investment

Managers ' investment decisions based on the criteria of maximizing shareholder value are considered to be efficient investments of enterprises, and vice versa. It is mainly manifested as over-investment or under-investment. Existing research shows that agency problems, information asymmetry, institutional hedging[11], and financing constraints, will promote the inefficient investment of enterprises. At present, data assets are still a new type of accounting element, and their relevant recognition rules, value measurement, and disclosure standards are not perfect enough [12][13]. For example, in the measurement of data assets, data assets classified as intangible assets may make the confirmation of costs more complicated due to strict capitalization conditions in the process of capitalization. The amortization or depreciation methods do not match the data assets and make them more volatile and more frequent in value fluctuations[9] Xu Xianchun et al. [14]also pointed out that China has not yet established a perfect database accounting method and system. Therefore, the accounting methods and statistical investigation systems for self-produced and self-used software and R&D activities still need to be further supplemented and improved.

On the other hand, with the rise of the digital economy, the process of enterprise digitization has accelerated, and the company has broadened its business scope and organizational structure based on the original, making the operational business boundary diversified and complex[15], which will lead to the intensification of the information asymmetry situation, making it more difficult for investors and regulators to supervise, and finally exacerbate the behavior of surplus management, leading to a reduction in investment efficiency. At the same time, the relevant accounting standards for data assets are currently not perfect enough, and it is highly likely to be flexibly manipulated by the management to achieve the real surplus management purpose. Wang Xinyi [9] has pointed out that when data assets are used as intangible assets, there are still many problems in the recognition of R&D expenses

DOI: 10.54254/2754-1169/104/20242686

in the process of association; BERESKIN et al[16] found that the management will cut R&D expenses because of the intention of surplus management, reduce the efficiency of research and development and then reduce the competitiveness of the company, which is also likely to affect the investment efficiency of the company.

To summarize, under the current imperfect system of data assets, data assets may cause investors to blindly pursue and increase the information asymmetry, financing constraints, and real surplus management problems, thus reducing the investment efficiency of enterprises. Therefore, this paper proposes the following hypotheses:

H1: Under other conditions, data assets will reduce enterprise investment efficiency.

2.2. Data Assets Enhance the Efficiency of Business Investment

The disclosure of data assets can make accounting information of higher quality and more comparable to reduce information asymmetry, and can also alleviate the agency problem to a certain extent[17]and reduce the situation of overinvestment. According to Biddle et al.[18], high-quality accounting information can reduce the risk of underinvestment and overinvestment, and companies that disclose high-quality accounting information are less affected by macroeconomic fluctuations. At the same time Yuan Zhenchao and Rao Pingui[19]believe that when allocating investment resources, the management can only make a reasonable comparative analysis to identify the risk and value of the investment if it fully understands the current situation of the enterprise and its situation in the industry, and this comparative analysis must be based on the basis that the accounting information between the enterprise and the middle of the enterprise is comparable. Therefore, with the continuous development of data assets, so that the quality of accounting information continues to improve, the market information barriers have been gradually eliminated, enterprises to obtain investment-related information more efficiently, to reduce the asymmetry of information as well as the cost of searching for information, thereby improving investment efficiency.

The disclosure of data assets allows managers to rely more on quantitative analysis of data for investment [20], which can increase the shareholders or investors to more comprehensive control of corporate information to reduce the probability of agency problems and make the enterprise more scientific and rigorous investment decisions, to improve the efficiency of the enterprise's investment. With the disclosure of enterprise data assets, the development of the digitalization process, when faced with financing problems, the extensive use of digital information makes banks and other financial institutions a more reasonable and comprehensive assessment of the enterprise's credit rating, especially for some small and medium-sized enterprises to come to[21], can make the credit enhancement [22], can be to a certain extent to alleviate the financing constraints, improve the enterprise's investment efficiency. extent alleviate the financing constraints and improve the financing efficiency of enterprises, which in turn can promote the efficient investment of enterprises.

In summary, data assets may be able to promote efficient investment by enterprises by alleviating their information disclosure asymmetry, agency problems, and financing constraints. Therefore, this paper proposes the following hypothesis:

H2: Data assets will enhance enterprises' investment efficiency under certain other conditions.

3. Research Design

3.1. Sample and Data Source

This paper selects Chinese A-share listed companies from 2013-2022 as the research object, excluding: (1) ST*, ST, and PT companies. (2) Listed companies in the financial industry. (3) Companies with serious missing data. The sources of data mainly include the Cathay Pacific (CSMAR) database, the Wind database, and annual reports of listed companies. To ensure the

DOI: 10.54254/2754-1169/104/20242686

accuracy of the data, the continuous variables in this paper have been shrink-tailed by 1% up and down. Through the above processing, a total of 13,180 balanced panel data are obtained.

3.2. Model Data and Variable Definitions

To test the hypotheses described in the previous section, the following baseline model is constructed in this paper:

$$Invest_t = \alpha_0 + \alpha_1 Data\_Asset + \alpha_2 Size + \alpha_3 Growth + \alpha_4 ROA + \alpha_5 Tobinq + \alpha_6 Cash + \alpha_7 Tang + \alpha_8 Lnage + \alpha_9 Indep + \alpha_{10} Fhold + \sum Year + \sum Industry + \varepsilon \quad (model1)$$

Invest represents the firm's investment efficiency, including overinvest and underinvest, and Data_Asset represents the value of the firm's data assets. Considering the possible impact of industry and year factors on the regression results, the model also controls for industry-fixed effects and year-fixed effects in addition to the basic control variables.

3.2.1. Measurement of Investment Efficiency

At present, most domestic scholars adopt Richardson's (2006) model to measure the investment efficiency of enterprises[23], and this paper also adopts this model to measure the investment efficiency of enterprises, and the specific model is shown in model 2:

$$Inv_t = \alpha_0 + \alpha_1 Growth_{t-1} + \alpha_2 Lev_{t-1} + \alpha_3 Cash_{t-1} + \alpha_4 Age_{t-1} + \alpha_5 Size_{t-1} + \alpha_6 Ret_{t-1} + \alpha_7 Inv_{t-1} + \sum Industry + \sum Year + \varepsilon \quad (model\ 2)$$

Among them, Invt represents the actual new investment level of the enterprise in year t, Invt-1 represents the amount of new capital investment in year t-1, and the residual represents the degree of deviation between the expected investment level and the actual investment level, which indicates the inefficient investment of the enterprise. The absolute value of the residual represents the investment efficiency of the enterprise, when the residual is greater than 0, it indicates that the enterprise overinvests, when the residual is less than 0, it indicates that the enterprise underinvests. $Growth_{t-1}$、Lev_{t-1}、$Cash_{t-1}$、Age_{t-1}、$Size_{t-1}$、Ret_{t-1} denote the growth rate of revenue, gearing ratio, cash holdings, number of years on the market, firm size, and annual stock return considering reinvestment of cash dividends in the year t-1, all of which are control variables; the model also controls for firms and years.

Table 1: Definition of Variables

Variables	Variable names	Description
Explained variables	Investment efficiency (INVEFF_abs)	The regression residuals are taken as absolute values
	Overinvestment(OverInv)	Regression residuals greater than 0
	Underinvestment(UnderInv)	Regression residuals less than 0 are taken as absolute values
Explanatory variables	Data_Asset	Ln(Market value-fixed assets-financial assets-intangible assets)
Control variables	Enterprise size(Size)	Number of employees
	Growth	growth rate of total assets
	Return on Assets(ROA)	The ratio of net profit to total assets
	Tobin's q ratio(Tobinq)	The ratio of market value to book value of the company
	Cash level(Cash)	Monetary funds / total assets
	Tangibility of assets(Tang)	(Total assets-net intangible assets) /total assets

DOI: 10.54254/2754-1169/104/20242686

Table 1: (continued).

Controlled Industries and Annual Fixed Efficiency	Number of years listed(Lnage)	Natural logarithm of the company's listing years
	Ratio of sole director(Indep)	The ratio of independent directors to the total number of directors
	Ownership concentration (Fhold)	The shareholding ratio of the largest shareholder (percentage)
	Equity nature (Soe)	The dummy variable is 1 for state-owned enterprises, otherwise 0.
	The life cycle of enterprise (lifecycle)	Virtual Variables, Ranking and Assigning Definitions
	Year	Annual dummy variables, control macroeconomic environment differences
	Industry(Ind)	Control industry differences

3.2.2. Measurement of Data Assets

At present, the measurement index standard of data assets is not perfect, and the existing measurement methods are based on the theoretical model derived from the use of "ln (market value-fixed assets-financial assets-intangible assets)" to measure[1]and based on the annual reports of listed companies disclosed in the details of the expenditure on fixed assets and intangible assets, according to the text mining to calculate the relevant frequency of the keywords and using the entropy method to weight of the index, and finally obtain the data assets index[2].

Because the first measurement method utilizes the income method of data asset value measurement, the difference between the market value of the enterprise and fixed assets, financial assets, and intangible assets is used to measure the value of data assets, which is more convincing for data-driven enterprises. Therefore, this measure is used as the core explanatory variable in this paper.

3.2.3. Control Variables

Based on the existing research literature on corporate investment efficiency[23][24], this paper selects firm size, growth (Tobinq), profitability (ROA), cash level (Cash), tangibility of assets (Tang), number of years on the stock market (Lnage), proportion of sole directors (Indep), equity concentration (Soe), and life cycle of the enterprise (lifecycle) as the basic control variables in this paper. In addition, we also control for year and industry-fixed effects.

3.3. Descriptive Statistics

The descriptive statistics of this paper are detailed in Table 2. For investment efficiency, the means of inefficient investment (INVEFF_abs), overinvestment (OverInv), and underinvestment (UnderInv) are 0.021, 0.029, and 0.017 respectively, and the mean of overinvestment is higher than that of underinvestment. The standard deviation of Overinvestment and Underinvestment are 0.033 and 0.014 respectively, and the median is 0.016 and 0.014 respectively, which shows that overinvestment is more serious than underinvestment and there is a bigger difference.

DOI: 10.54254/2754-1169/104/20242686

Table 2: Descriptive Statistics of Variables

Variables	N	Mean	Max	SD	p75	p50	p25	Min
INVEFF_abs	13180	0.021	0.150	0.023	0.027	0.015	0.007	0.000
OverInv	5130	0.029	0.178	0.033	0.039	0.016	0.006	0.000
UnderInv	8050	0.017	0.097	0.014	0.023	0.014	0.007	0.000
Data_Asset1	13180	23.06	26.66	1.10	23.74	22.89	22.25	20.79
Size	13180	6700	79000	12000	6400	2900	1400	75
Growth	13180	0.141	4.370	0.370	0.228	0.090	-0.027	-0.682
ROA	13180	0.036	0.229	0.057	0.062	0.033	0.013	-0.680
Tobinq	13180	1.984	15.607	1.297	2.288	1.574	1.184	0.814
Cash	13180	0.145	0.662	0.113	0.188	0.113	0.067	0.007
Tang	13180	0.927	1.000	0.084	0.977	0.954	0.913	0.452
Lnage	13180	2.914	3.434	0.327	3.258	2.944	2.565	2.485
Indep	13180	0.376	0.600	0.056	0.429	0.364	0.333	0.308
Fhold	13180	33.73	75.46	14.76	44.02	31.77	21.90	7.52

For Data Asset, the mean is 23.063, the standard deviation is 1.101, the median is 22.887, the P75 is 23.737, and the P25 is 22.251, which indicates that the average level of Data Asset is in the range of 22-23. In addition, the maximum value of data assets is 26.663 and the minimum value is 20.789, which indicates that the value of data assets in the sample varies greatly. The descriptive statistics of the other control variables are consistent with other existing literature[23][24].

4. Empirical Testing

4.1. Benchmark Regression Analysis

The results of the benchmark regression analysis of data assets and corporate investment efficiency are shown in Table 3. The study found that in the full sample and over-investment samples, data assets and corporate investment efficiency are significantly positively correlated at the 1 % level; data assets have a significant role in promoting the inefficient investment of enterprises, especially the over-investment of enterprises. The results support Hypothesis 1.

In the economic sense, when the data assets increase by one unit for all samples, it will cause INVEFF _ abs to increase by 0.004. accounting for 19.04 % of the average INVEFF_abs. For the over-investment sub-sample, when the data assets increase by one unit, OverInv will increase by 0.008, accounting for OverInv. 27.59 % of the value. For the under-investment sub-sample, when the data asset increases by one unit, it will lead to an increase of 0.001 in UnderInv, accounting for 5.88 % of the average UnderInv.

Whether in a statistical or economic sense, the rise of data assets will significantly reduce the investment efficiency of enterprises, especially the over-investment of enterprises.

DOI: 10.54254/2754-1169/104/20242686

Table 3: Benchmark Test Results

Variable	(1) INVEF_abs	(2) OverInv	(3) UnderInv
Data_Asset1	0.004***	0.008***	0.001*
	(5.44)	(4.97)	(2.05)
Size	-0.000*	-0.000*	0.000
	(-1.99)	(-2.34)	(0.21)
Growth	0.001*	0.002	0.000
	(2.36)	(1.73)	(0.91)
ROA	0.006	0.010	0.000
	(1.45)	(0.76)	(0.04)
Tobinq	-0.001**	-0.002	-0.000
	(-2.68)	(-1.91)	(-1.27)
Cash	0.003	0.004	0.001
	(1. 11)	(0.60)	(0.33)
Tang	-0.004	-0.020	-0.003
	(-0.76)	(-1.46)	(-0.77)
Lnage	-0.301*	-0.564**	0.104
	(-2.26)	(-3. 14)	(1.57)
Fhold	0.000	0.000*	0.000
	(1.83)	(1.98)	(1.43)
_cons	0.822*	1.524**	-0.281
	(2. 12)	(2.88)	(-1.46)
Ind FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
N	13180	5130	8050
R2	0.038	0.047	0.062

t statistics in parentheses
* p < 0.05, ** p < 0.01, *** p < 0.001

Table 4: gradually increasing control variables.

Variable	(1)	(2)	(3)	(4)
Data_Asset1	-0.001**	0.003***	0.004***	0.004***
	(-2.78)	(5.62)	(5.57)	(5.44)
Size			-0.000	-0.000*
			(-1.89)	(-1.99)
Growth			0.002*	0.001*
			(2.52)	(2.36)
ROA			0.007	0.006
			(1.82)	(1.45)
Tobinq			-0.001**	-0.001**
			(-2.81)	(-2.68)
Cash				0.003
				(1. 11)
Tang				-0.004
				(-0.76)
Lnage				-0.301*
				(-2.26)
Fhold				0.000
				(1.83)
_cons	0.033***	-0.038*	-0.053*	0.822*
	(7.80)	(-2.01)	(-2.53)	(2. 12)
Ind FE	No	Yes	Yes	Yes
Year FE	No	Yes	Yes	Yes
N	13180	13180	13180	13180
R2	0.001	0.034	0.037	0.038

t statistics in parentheses
* p < 0.05, ** p < 0.01, *** p < 0.001

4.2. Robustness Test

4.2.1. Multicollinearity Test

To avoid the high correlation between the independent variables in the model, the estimation of the regression model is unstable, so the Variance Inflation Factor (VIF) index is used for testing. The VIF values of the independent variables in this paper are all less than 3, indicating no high degree of multicollinearity in the model, and the model is more robust.

Table 5: VIF Test Results

Variable	VIF	1/VIF
Data_Asset1	1.73	0.58
Size	1.66	0.60
ROA	1.20	0.83
Tobinq	1.20	0.83
Cash	1.11	0.90
Lnage	1.09	0.92
Fhold	1.08	0.92
Growth	1.08	0.93
Tang	1.05	0.95

DOI: 10.54254/2754-1169/104/20242686

Table 5: (continued).

Indep	1.02	0.98
Mean VIF	1.22	

4.2.2. Variable Substitution—Replacement of Explanatory Variables

In the robustness test, drawing on the practice, the new data asset index is recorded as Data _ Asset2[25], which includes monetary funds, purchase and resale financial assets, dividends receivable, interest receivable, short-term investment, long-term equity investment, and long-term receivables. After replacing the explanatory variables, the regression of inefficient investment (INVEFF_abs), overinvestment, and underinvestment are carried out again.

The result is that Data _ Asset2 has a positive impact on inefficient investment, overinvestment, and underinvestment, and both inefficient investment and overinvestment are significant at the 1 % level. Further considering the total assets recorded in the company's financial statements, the part of the company's market value that exceeds the total assets is used as an alternative indicator of the company's data assets, recorded as Data_Asset3. By replacing the explanatory variables again, the result is still that Data _ Asset3 has a significant positive impact on the overall inefficient investment and overinvestment of the enterprise. This shows that after replacing the measurement method of explanatory variables, the previous conclusions are not changed, indicating that the basic regression results are robust.

Table 6: Substitute the Robustness Test of Explanatory Variables

Variable	(1) INVEFF_abs	(2) OverInv	(3) UnderInv	(4) INVEFF_abs	(5) OverInv	(6) UnderInv
Data_Asset2	0.004***	0.007***	0.001			
	(5.57)	(4.92)	(1.87)			
Data_Asset3				0.003***	0.008***	0.000
				(3.81)	(4.22)	(0.35)
Control variable	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	13180	5130	8050	13180	5130	8050
R2	0.038	0.047	0.062	0.036	0.045	0.061

t statistics in parentheses
* p < 0.05, ** p < 0.01, *** p < 0.001

4.2.3. Quantile Regression

To explore whether data assets and non-efficiency investment of enterprises are affected by the level of non-investment efficiency of enterprises, this paper makes regressions at the 20th, 40th, 60th, and 80th quantiles of non-investment efficiency respectively, and also avoids the defects of Richardson's model applicable to developed capital markets. The regression results are shown in table 7. The impact of data assets on the inefficient investment of enterprises is positive and significant when the inefficient investment of enterprises is low. When the non-investment efficiency of enterprises is high, the positive impact of data assets on the non-investment efficiency of enterprises is not significant. It shows that the non-efficiency level of the enterprise itself is low, and the role of more data assets is

DOI: 10.54254/2754-1169/104/20242686

more likely to cause the decline of the enterprise's investment efficiency. The test results of the benchmark regression are verified again.

Table 7: Shows the Robustness Test of Quantile Regression

Variable	(1)20 INVEFF_abs	(2)40 INVEFF_abs	(3)60 INVEFF_abs	(4)80 INVEFF_abs
Data_Asset1	0.002* (2. 16)	0.003** (3.20)	0.004** (2.62)	0.006 (1.91)
Control variable	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
N	13180	13180	13180	13180

t statistics in parentheses
* p < 0.05, ** p < 0.01, *** p < 0.001

4.2.4. Endogeneity Problem Treatment for Sample Bias - Heckman Phase II

Although the degree of data assets can be observed by all enterprises, some enterprises do not invest, and only enterprises with investment behavior can be observed. Therefore, some investment efficiencies cannot be observed in the benchmark model of this paper or not measured in the Richardson model. To better control the problem of sample deviation, this paper chooses Heckman's two-stage model to deal with it. In the first stage of the model, this paper sets the dummy variable according to whether the data assets of the enterprise exceed the industry average. If it exceeds the industry average, the value is assigned to 1, and if it does not exceed, the value is assigned to 0. The inverse Mills ratio (IMR)[2]was calculated using the Probit model. In the second stage, IMR is added for regression to observe the influence degree of explanatory variables. It can be seen from Table 8 that the impact of data assets on corporate inefficient investment is still significant at the 1 % level after eliminating sample selection bias.

Table 8: The Endogenous Test Results of Sample Deviation

Variable	(1)Dataif	(2)INVEFF_abs
Data_Asset1	2.022*** (55.94)	0.004*** (5.24)
Size	0.000*** (12.84)	-0.000** (-2.83)
Growth	-0.000 (-0.00)	0.001** (2.70)
ROA	2.741*** (7.85)	0.006 (1.41)
Tobinq	0.015 (1.05)	-0.001*** (-3.54)
Cash	0.366* (2.37)	0.003 (1.31)
Tang	-0.481* (-2.48)	-0.004 (-0.95)
Lnage	-0.313*** (-5.86)	-0.301** (-2.64)
Fhold	-0.001 (-0.78)	0.000* (2.28)
IMR1		0.000 (0.33)
_cons	-45.460*** (-54.70)	0.817* (2.46)
Industry FE	Yes	Yes
Year FE	Yes	Yes
N	13180	13180
R2		0.038

t statistics in parentheses
* p < 0.05, ** p < 0.01, *** p < 0.001

DOI: 10.54254/2754-1169/104/20242686

4.2.5. Omission Variables and Reverse Causal Endogeneity Problem-Instrumental Variable Method

To avoid omitted variables and reverse causation problems in the model, this paper adopts a static GMM model.

Table 9: Results of Endogeneity Tests for Omitted Variables and Reverse Causation

Variable	(1)Data_Asset1	(2)INVEFF_abs
IV	0.733*** (8. 15)	
L.IV	0.201* (2.20)	
(Kleibergen-Paap rk LM statistic):	132.349	
(Kleibergen-Paap rk Wald F statistic):	77.302 [19.93]	
Data_Asset1		0.0115*** (13.01)
N	11857	10539
F	440.4	169.3

Hansen J statistic (overidentification test of all instruments): 2.538
Chi-sq(1) P-val =0.1112
t statistics in parentheses
* p < 0.05, ** p < 0.01, *** p < 0.001

First,the instrumental variable (IV) is set to be the average of the data assets of other firms in the firm's industry and lagged by one period. In the first stage, the Kleibergen- Paaprk LM statistic and Kleibergen-Paaprk Wald F statistic indicate the instrumental variable passes the under-identification test and weak identification test, and the Hansen test is not significant. Hansen's test is not significant, representing the rationality of the model. The instrumental variable (IV) and the one period lagged instrumental variable (L.IV) are significantly and positively correlated with data assets, indicating that the degree of ownership of data assets by other firms in the industry has a positive impact on the firm's data assets, which is related to the strong liquidity of data assets. In the second stage, data assets still have a significant positive effect on firms' inefficient investment, and the conclusion that data assets can promote firms' inefficient investment remains unchanged, as in the previous section.

4.3. Heterogeneity Analysis

4.3.1. Heterogeneity Analysis of Different Industries

Enterprises with different property rights have significant differences in the management level of data assets, as well as investment preferences and investment behavior[26]. Therefore, this paper will deduce the different degrees of impact of data assets on investment efficiency in various industries, as shown in Table 10.

Table 10: Heterogeneity Analysis of Different Equity Natures

Variable	(1)	(2)	(3)	(4)	(5)	(6)
	INVEFF_abs	INVEFF_abs	OverInv	OverInv	UnderInv	UnderInv
	state-owned	non	state-owned	non	state-owned	non
Data_Asset1	0.004*** (3.69)	0.005*** (4.69)	0.006* (2.52)	0.012*** (4.52)	0.002* (2.21)	0.001 (1.58)
Controls	Yes	Yes	Yes	Yes	Yes	Yes

DOI: 10.54254/2754-1169/104/20242686

Table 10: (continued).

Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	5882	7298	2324	2806	3558	4492
R2	0.038	0.045	0.059	0.055	0.064	0.075

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In this paper, enterprises are divided into state-owned enterprises and non-state-owned enterprises according to their ownership. State-owned enterprises are assigned the value of 1, and non-state-owned enterprises are assigned the value of 0. As shown in columns (1) and (2) of Table 10, data assets can significantly affect the inefficient investment of firms, and it is more significant in non-state-owned firms. This suggests that data assets reduce the investment efficiency of non-state-owned enterprises more than large state-owned enterprises, which may be related to the serious problem of financing constraints of small and medium-sized private enterprises, and non-state-owned enterprises may reduce their investment efficiency more because of the problem of financing difficulties. Columns (3) and (4) show that data assets are more effective in promoting the investment efficiency of non-state enterprises than state-owned enterprises. Columns (3) and (4) suggest that data assets are more likely to contribute to overinvestment in non-state firms than in state-owned firms, which may be related to the nature of non-state firms, which mostly focus on innovation, which may lead to overinvestment behavior. Columns (5) and (6) show that data assets are less useful for underinvestment in non-state firms than in state-owned firms, which further confirms the above conclusion.

4.3.2. Heterogeneity Analysis of Enterprise Life Cycle

Enterprises in different life cycle stages are affected by different external environments and therefore have different internal resource conditions, which leads to different investment and financing efficiency of enterprises, and differences in investment decisions and investment efficiency. While listed enterprises pass through the start-up period, therefore, the life cycle of enterprises is mainly divided into the following stages: growth period, maturity period, and decline period[27].

According to Li Yunhe's method, the sales revenue growth rate, retained earnings rate, capital expenditure rate, and age of enterprises were ranked and assigned values, and the sample data were sorted from largest to smallest, and the samples of each industry were divided into three parts according to the size of the total scores, and the one-third of the part with the highest score was defined as the growth stage enterprises, the one-third of the part with the lowest score was defined as the decline stage enterprises and the middle part was defined as the maturity stage enterprises. The middle portion is defined as mature firms.

As shown in columns (1), (2), and (3) of Table 11, data assets contribute more to the inefficient investment of growing and declining firms, i.e., an increase in data assets is more likely to lead to a decrease in the investment efficiency of firms located in growing and declining periods. This is even more obvious in the case of overinvestment by data assets. Firms located in the growth period generally have stronger investment preferences[28], so the increase in data assets is more likely to lead to overinvestment behavior, while firms located in the decline period may have difficulty in coping with and managing overinvestment.

DOI: 10.54254/2754-1169/104/20242686

Table 11: Heterogeneity Analysis across Life Cycles

	(1)	(2)	(3)	(4)	(5)	(6)
Variable	Growth period	Maturity period	Decline period	Growth period	Maturity period	Decline period
	INVEFF abs	INVEFFabs	INVEFF abs	OverInv	OverInv	OverInv
Data_Asset1	0.010***	0.005**	0.004***	0.022***	0.006	0.008***
	(4.04)	(2.86)	(3.83)	(4.85)	(1.48)	(3.73)
N	2703	4304	6162	1429	1802	1895
R2	0.057	0.035	0.036	0.095	0.045	0.086

t statistics in parentheses
* p < 0.05, ** p < 0.01, *** p < 0.001

5. Mechanism Test

In the previous study, this paper has obtained the conclusion that data assets can significantly reduce the investment efficiency of enterprises. Under the current premise that data assets disclosure is not standardized, which may cause blind pursuit by investors and lead to data bubbles. Therefore, the higher the level of data assets disclosed by enterprises will have a certain adverse effect. The specific mechanism will be explored through the two aspects of enterprise financing constraints and real earnings management.

5.1. Corporate Financing Constraints

In the current context, more data assets will increase the problem of financing constraints. As shown in Table 12, data assets have a significant positive impact on financing constraints (SA), which indicates that the higher the degree of corporate data assets, the more financing constraints. A large number of scholars have proved that the increase in financing constraints will reduce the investment efficiency of enterprises[29][30][31][32].

Table 12: Mechanistic Tests of Corporate Financing Constraints

Variable	(1) SA	(2) SA
Data_Asset1	0.031***	0.018***
	(9.48)	(4.84)
Control variable	No	Yes
Industry FE	Yes	Yes
Year FE	Yes	Yes
N	12960	12960
R2	0.858	0.883

t statistics in parentheses
* p < 0.05, ** p < 0.01, *** p < 0.001

First of all, as mentioned in the previous article, there are still many imperfections in the measurement and disclosure of data assets.

In the face of uncertainty, if enterprises blindly increase data assets, it will increase information asymmetry. Because of the increase in the extent of data assets, internal shareholders and management have a better understanding of the company's situation, but at the same time, external investors have a lower degree of understanding of the true information of the enterprise, and are unable to understand the true situation of the company because of the greater uncertainty. As

DOI: 10.54254/2754-1169/104/20242686

information asymmetry arises, the external cost of capital for the firm becomes higher, as each investor does not have full and accurate information about the firm's current situation, leading to more uncertainty, and a higher risk premium is required by the investors, which increased the external financing constraints.

Secondly, in the current context, the increase in the extent of data assets may increase the creation of agency problems. Based on the assumption of "economic man", managers will be more inclined to act and make decisions based on the principle of profit maximization. Therefore, when faced with an increase in the extent of data assets, it becomes more difficult for managers to manage, and the functions and qualities required of them increase. The agency problem deepens as managers with extensive expertise often choose to sacrifice the interests of the group to maximize their gain. When the agency problem is more serious, the external investors perceive that their interests may be infringed upon by insiders, which can lead to financing constraints.

5.2. True Corporate Surplus Management

In the current context, more data assets will promote the enterprise's real earnings management. As can be seen in Table 13, data assets have a significant positive impact on the real earnings management of enterprises, which indicates that the higher the degree of data assets of enterprises, the higher the degree of real earnings management. A large number of scholars have proved that the real earnings management of enterprises will lead to a decline in the investment efficiency of enterprises[33][34][35].

Table 13: Mechanism Tests for Real Earnings Management by Firms

Variable	(1) RM	(2) RM
Data_Asset1	-0.029***	0.010*
	(-7.57)	(2. 17)
Control variable	No	Yes
Industry FE	Yes	Yes
Year FE	Yes	Yes
N	13060	13060
R2	0.025	0.070

t statistics in parentheses
* p < 0.05, ** p < 0.01, *** p < 0.001

As mentioned earlier, when the recognition of R&D expenses is not clear when the data asset is an intangible asset, management may cut or increase R&D expenses to achieve the purpose of real earnings management. With the capitalization of data, management can seize the loophole that the measurement criteria for data assets are not clear and sell the assets in such a way that there is a big difference between the book value and the market value of the data assets to further adjust the profit and loss of the enterprise.

In this way, the difference between the book value and the market value of the data assets can be further adjusted to achieve the purpose of real earnings management. In addition, the agency problem caused by data assets can also lead to the emergence of real earnings management. In the existing literature, it has been shown that to meet analysts' surplus expectations, management has the incentive to manage earnings to reduce the agency problem with shareholders[36], which suggests that in the current context, the deepening of agency problems caused by data assets may prompt management to resort to real earnings management in an attempt to reduce it, and thus the extent of real earnings management by the firm may increase.

DOI: 10.54254/2754-1169/104/20242686

6. Conclusions and Implications of the Study

With the trend of data assets in the financial statements, how to deeply study the situation of data assets at the present stage and the impact, to obtain the future data assets for the enterprise to bring more stable development of the enlightenment for the research objectives of this paper. In the case of imperfect data asset measurement standards, this paper quantifies the extent of data assets of listed companies concerning the data asset measurement methods that other scholars have already unearthed so far and studies the impact of data assets on the investment efficiency of enterprises, and the conclusions remain unchanged through a series of robustness tests; and finally, through the mechanism test, we find that it is the imperfect measurement and disclosure of the existing data assets that data assets will through increasing the financing constraints of enterprises and the real earnings management of enterprises and then lead to the decline of their investment efficiency. This provides some ideas for our understanding of data assets. Therefore, the findings of this paper have the following insights:

(1) Enterprises and their stakeholders should take a dialectical view of data assets. Data assets can have a range of positive impacts on an enterprise, but only if the measurement and disclosure of data assets have been more maturely developed. In the current situation where data assets are still on their way to entering the financial statement, enterprises should handle data assets with caution. When expanding the scale of data assets, the value of data assets should be explored in depth, to prevent data assets from being overvalued, which may lead to blind investment and lower investment efficiency.

(2) Non-state-owned enterprises and enterprises in growth and decline should pay more attention to the development of data assets, as it has been confirmed in the previous article that due to the serious financing problems of non-state-owned enterprises, data assets are more capable of promoting over-investment in non-state-owned enterprises than in state-owned enterprises. The increase in the degree of data assets of non-state enterprises will lead to an increase in their investment in innovation, but enterprises should not invest blindly because of this, and they should make good use of their data assets by managing and applying them well while carrying out innovation. Enterprises in the growth and decline periods should also pay attention to the fact that the increase in data assets will allow them to invest blindly to utilize sufficient free cash flow and to cope with the risk of mismanagement.

(3) As far as external regulation is concerned, national regulators should further improve laws and accounting standards accounting standards on data assets, and clear and authoritative measurement standards for data assets to cope with the complicated data in the era of big data. The government should pay attention to the role played by data assets in the economy and promote the transformation and upgrading of traditional manufacturing industries by strictly regulating data assets while preventing the abuse of data assets and the emergence of data bubbles, to provide good support for enterprises in the development of the digital economy.

Although this paper has made some research and conclusions on the impact of data assets on enterprise investment efficiency, the current standard of measuring enterprise data assets is still not perfect, so the conclusions of the study have some limitations. How to further measure the value of enterprise data assets accurately and classify and analyze the impact of data assets on various types of enterprises is the focus of subsequent research.

Authorship Contribution Statement

First author: searching for relevant literature, proposing and constructing the writing framework; using “Stata” to conduct empirical research on the data; and drawing experimental conclusions after

DOI: 10.54254/2754-1169/104/20242686

multiple tests and analyses of the data results: proposing relevant suggestions and insights; paper layout.

Corresponding author: searching for relevant literature, proposing and discussing writing ideas; collecting the data required for the study; assisting in completing the empirical study and related analyses; completing the literature review, theoretical foundations, and hypotheses; translating and submitting the paper.

Second author: Participate in the discussion of the writing framework; collect the data required for the empirical evidence; assist in analyzing the results; translate the paper, polish the language, and layout the paper.

References

[1] *Lu,Z., Zhou,T., Wang, L., Liao, Z.,2023. Data Assets and Corporate Development: Empirical Evidence from Chinese Listed Companies[J]. Industrial Economics Research. (04):128-142.*

[2] *Li,J., Dong,X., Zhang,J., Tao,Y.,2023.Research on the impact of data assets on corporate innovation investment[J]. Foreign Economy and Management.45(12):18-33.*

[3] *Lu,M., Wang,W., Zhu,Z.,2023. How Data Assets Enable High-Quality Development of Enterprises: Empirical Evidence from the Perspective of Capacity Utilization[J]. Journal of Shanghai Business School,24(04):22-41.*

[4] *Charles,I.,Jones,C.T.,2020.Nonrivalry and the Economics of Data.AMERICAN ECONOMIC REVIEW,110(9):pp. 2819-58.*

[5] *Hajli, Mahmood;Sims,Julian M.;Ibragimov, Valisher,2015.The international journal of productivity and performance management. Information technology (IT) productivity paradox in the 21st century.64 (4): 457 – 478.*

[6] *Liu,L.,Yang,G.,2022.Research on the Impact of Digital Construction on Corporate Investment Efficiency--Empirical Evidence from Listed Companies in China's Manufacturing Industry[J].Accounting Friends.(23):123-130.*

[7] *Liu,Y.,Tan,H., Chen,X., Yang,C.,2022.Research on the impact of digital economy development on investment efficiency of real economy[J].China Soft Science.(10):20-29.*

[8] *Zhou,Y.,Sun,H.,Liu,R.,2024.The impact of the digital economy on the investment efficiency of small and medium-sized enterprises in Daqing[J].Cooperative Economy and Science and Technology.(15):56-57.*

[9] *Wang,X.,2023.Influence and Countermeasures of Data Resource Accounting Processing Methods on Enterprise Financial Risk[J]. Small and medium-sized enterprise management and science and technology,(01):191-193.*

[10] *Li,X.,He,W.,2022.Digital Innovation, Business Environment and High-Quality Development of Enterprises: Empirical Evidence Based on New Generation IT Industry Listed Companies[J].Science and Technology Management.43(11):56-77.*

[11] *Jeremy C. S.,2003.Agency,Information,and Corporate Investment.Handbook of the Economics of Finance.1 (A): 111-165.*

[12] *Zhang,J.,Gu,Y.,Song,X.,2020.Research on accounting treatment and information presentation of enterprise data assets[J]. Accounting and Economic Research.34(03):3-15.*

[13] *Liu,Y.,2014.Recognition and Measurement of Big Data Assets [J]. Business Accounting.(18):3-4.*

[14] *Xu,X.,Zhang,Z.,Hu,Y.,2022.A study on statistics and accounting of data assets [J]. Management World,38(02):16-30+2.*

[15] *Li,R., Wang,Y., Lu,Z.,2020.Do internet business models affect the surplus quality of listed companies--Empirical evidence from Chinese stock market[J]. Accounting Research.(10):66-81.*

[16] *Frederick L. Bereskin, Po-Hsuan Hsu,Wendy Rotenberg,2018.The Real Effects of Real Earnings Management: Evidence from Innovation. Contemporary Accounting Research.35 (1): 525-557.*

[17] *Michael C. Jensen,1986.Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers. The American Economic Review.76 (2):323-329.*

[18] *Gary C. Biddle, Gilles Hilary, Rodrigo S. Verdi.,2009.How does financial reporting quality relate to investment efficiency? Journal of Accounting and Economics.48 (2–3): 112-131.*

[19] *Yuan,Z., Rao,P., 2018.Accountition comparability and investment efficiency[J]. Accounting Research.(06):39- 46.*

[20] *Li,L., Yang,S., Chen,N.,2022. Research on the impact of digital transformation on enterprise investment efficiency[J]. Soft Science.36(11):23-29.*

[21] *Tang,S.,Wu,X.,Zhu,J.,2020.Digital finance and corporate technological innovation - structural characteristics, mechanism identification and effect differences under financial regulation[J]. Management World.36(05):52-66+9.*

[22] *Chen,X.,Zhang,H.,2021.How the digital economy affects the level of risk-taking [J]. Economic Management.43(05):93-108.*

DOI: 10.54254/2754-1169/104/20242686

[23] Yao,L., Chen,X., Zhou,Y.,2020.Management Capability and Investment Efficiency [J]. Accounting Research.(04):100-118.
[24] Gao,J.,Chu,D.,Lian,Y.,Zheng,J.,2021.Can ESG performance improve corporate investment efficiency? [J].Securities Market Herald.(11):24-34+72.
[25] Hu,Y.,Wang,X.,Zhang,J.,2017.Motivation of financial asset allocation:"Reservoir"or"Substitution"?--Evidence from Chinese listed companies[J]. Economic Research.52(01):181-194.
[26] LI,Y.,Qin,Y.,Zhang,X.,2011.Corporate Property Rights, Managerial Background Characteristics and Investment Efficiency[J]. Management World.(01):135-144.
[27] Li,Y.,Li,Z.,Tang,S.,2011. Corporate life cycle, corporate governance and corporate capital allocation efficiency[J].Nankai Management Review. 14(03):110-121.
[28] Xie,P.,Wang,C.,2017.Management Power, Firm Life Cycle and Investment Efficiency-An Empirical Study Based on Listed Companies in China's Manufacturing Industry[J]. Nankai Management Review.20(01):57-66.
[29] Lian,Y.,Su,Z.,2009.Financing constraints, uncertainty and investment efficiency of listed companies[J].Management Review. 21(01):19-26.
[30] Guo,Q.,Luo,B.,2013.Financing constraints,accounting information quality,and investment efficiency[J].Journal of Zhongnan University of Finance and Law.(01):102-109.
[31] Yu,K.,Li,Z.,Zhang,X.,2014.The mystery of corporate investment efficiency: Financing constraint hypothesis and monetary policy impact[J]. Economic Research.49(05):106-120.
[32] Han,Y.,Zhao,Z.,Shi,B.,2021.Financial Development, Monetary Policy Changes and Corporate Investment Efficiency: An Empirical Study Based on the Perspective of Corporate Finance Constraints[J].Economic and Management Review.37(04):30- 43.
[33] Liu,H.,Wang,C.,Wu,L.,2014. Decision-making power allocation, surplus management, and investment efficiency[J]. Economic Research.49(08):93-106.
[34] Gu,M.,Yang,J., Yu N.,2016. The Impact of Surplus Management on Corporate Investment Decisions Study Based on Accrual Surplus Management and Real Surplus Management[J]. Contemporary Accounting Review.9(01):49-69.
[35] Luo,X.,Fu,Y.,2017. The impact of real surplus management on corporate inefficient investment[J]. Technology and Management.19(04):78-83.
[36] David Burgstahler ,Ilia Dichev,1997.Earnings Management to Avoid Earnings Decreases and Losses[J]. Journal of Accounting and Economics.(24): 99-126.