Empirical Analysis of Enterprise Digital Transformation Degree on Listed Manufacturing Companies' Performance from the Digital Economy Perspective

Kehan Lu*

School of Economics, Guangdong University of Technology, Guangzhou, 510006, China

Abstract. As a cornerstone of China's economy, the manufacturing sector's digital transformation is crucial for corporate sustainability and vital to the structural optimization of the national economy. Using panel data from China's manufacturing A-share listed firms (2014–2023), this paper empirically analyzes the impact of corporate digital transformation on business performance using text mining and econometric methods. It is found that the coefficient of digital transformation is 0.0063, which significantly enhances the financial performance of enterprises. The regression analysis yields a parameter estimate of 0.0086 for private sector entities, which contrasts with the negative coefficient for state-owned entities (-0.0031), suggesting differences in the strength of the ownership structure among them. This paper reveals the mechanism by which digital transformation promotes performance improvement by optimizing resource allocation and operational efficiency through fixed effects model and group regression analysis. At the same time, this study establishes a theoretical foundation for manufacturing firms' digital transformation and proposes differentiated policy measures: strengthen the precise policy inclination for non-state-owned enterprises (NSOEs); establish a system to assess the success of digital transformation in state-owned enterprises (SOEs); promote the construction of a cross-enterprise data-sharing platform; and foster collaborative progress in digital transformation throughout the industry.

1 Introduction

During the digital economy era, breakthroughs in new generations of information technology such as Artificial Intelligence, Blockchain, Cloud Computing, and Big Data (the “ABCD” technologies). The introduction of these technologies by an enterprise to redesign its value creation mechanisms and operational architecture to achieve competitive alignment with the market dynamics of the digital age is called enterprise digital transformation [1]. In the face of profound changes in technology and market environment, more and more Chinese

* Corresponding author: 3123008092@mail2.gdut.edu.cn

companies are using digital tools and platforms to drive organizational change and accelerate product and service innovation.

The Report released by the China Digital Economy Research Institute in 2024 shows that in terms of expansion, in 2023, China's digital economy expanded to 53.9 trillion yuan, marking an increase of 3.7 trillion yuan from the prior year. This growth rate has now stabilized within a more consistent range [2]. The growing importance of the digital economy within the national economy has become even more evident, especially regarding its contribution to efficiency. China's digital economy reaches 42.8% of GDP in 2023, up 1.3% from the previous year. According to the data in Figure 1, the scale of China's digital economy has increased from 11.2 trillion yuan in 2012 to 53.9 trillion yuan in 2023, and although the growth rate has slowed down in recent years, it has remained above 7%.

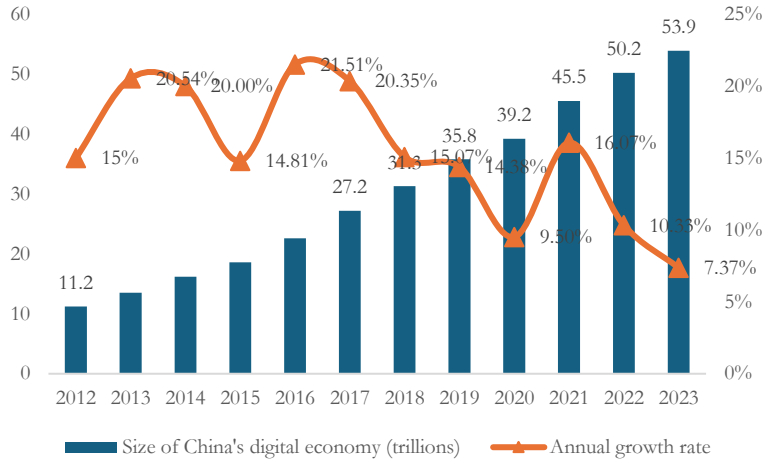

Fig. 1 Dynamic changes in market scale and its YOY growth rate of digital economy, China

Data source: China Academy of Information and Communications Technology, National Bureau of Statistics
Photo credit: Original

Following the landmark 20th CPC National Congress, strategic initiatives have been implemented nationwide to accelerate the integration of digital technologies within industrial production systems. The "Action Plan for the Digital Transformation of Manufacturing" formulated at the State Council Executive Meeting on May 11, 2024 emphasized that the digital transformation of manufacturing is an important measure to promote new industrialization and build a modern industrial system [3]. At the National People's Congress held on March 5, 2025, the government work report once again proposed that the digital transformation of manufacturing should be accelerated.

This paper uses text mining and Python's Jieba library to extract the frequency of keywords such as digital technology, big data, blockchain, cloud computing, and artificial intelligence, and measures the degree of digitalization of listed manufacturing companies. The specific calculation method is to use the application in the annual report of listed manufacturing companies as the keyword of digital transformation, and measure digital transformation with its word frequency [4]. Due to the prominent "right-biased" nature of word frequency data (most companies have a low degree of digitalization, but a few companies have advanced transformation), this paper uses the logarithmic processing method of adding 1 in the data processing. The logarithmic processing is used to reduce the maximum value in these data, thereby obtaining a more balanced and suitable overall indicator for

analysis, which is employed to assess the general state of enterprise digital transformation and subsequently derive a comprehensive metric that reflects this transformation.

This paper selects return on equity (ROE) as a measure of corporate operating performance, takes the operating performance of China's A-share listed manufacturing companies from 2014 to 2023 as the research object, establishes an econometric model, proposes hypotheses H1 and H2, and uses STATA to conduct empirical analysis.

2 Methods

2.1 Data sources

The financial information used in this study is sourced from the CSMAR database, which provides detailed financial data for analysis, and the sample data is processed as follows: 1. Eliminate samples from non-manufacturing industries; 2. Eliminate samples that have been specially treated (ST) or delisting risk warnings (*ST) during the entire sample period; 3. Eliminate samples with a listing time of less than one year; 4. Eliminate samples that have been suspended or terminated during the sample period; 5. Remove data entries where values are absent for a period exceeding five consecutive years. For variables with gaps in data spanning fewer than five years,hierarchical interpolation is implemented: trend adjustment interpolation is applied to time series variables, and multiple interpolation is used for key indicators [5, 6]. Therefore, this paper finally obtained 9168 observation data of 1236 enterprises from 2014 to 2023, which are balanced panel data.

2.2 Variable selection and processing

This paper chooses business performance as the dependent variable and specifically applies the return on equity (ROE) in the annual reports of enterprises to measure business performance.

This paper selects digital transformation (Digital) as the dependent variable. Currently, most research on digital transformation focuses on qualitative analysis, with relatively few quantitative studies. Existing studies have used dummy variables (such as whether or not digital transformation has been carried out) for measurement, but this method cannot accurately reflect the "intensity" of the transformation. Therefore, the frequency of keywords in the annual reports of listed companies is analyzed as a proxy for measuring the digital transformation of companies [4]. In this study, Python-based web scraping tools are employed to gather annual reports from all A-share listed companies on both the Shanghai and Shenzhen stock markets. The text data is then processed using the Java PDFBox library to extract relevant content, facilitating the creation of a digital transformation keyword database. The keyword library was developed based on academic research and industry policy documents that were screened and categorized. By counting the frequency of words related to digital transformation in the annual report; in order to reduce data bias, logarithmic processing is used to shrink the data by 1% up and down to form the indicator system of enterprise digital transformation. To ensure the validity of the keyword library, this paper randomly extracted 100 annual reports and manually labeled the digital transformation-related content, and compared with the algorithmic extraction results, the accuracy rate reached 89.3%.

Referring to the existing literature [1, 4, 7], this paper selects company size (Size); leverage ratio (Lev); employee density (StaffIntensity); average age of management (AverageAge); price-to-book ratio (PB); years of establishment (FirmAge); proportion of

intangible assets (Intangible); financial leverage (FL); Tobin's Q value (TobinQ). These 10 variables are used to control the model.

All control variables were assessed for multicollinearity using the Variance Inflation Factor (VIF), with the findings presented in Table 1:

Table 1 Multicollinearity between control variables

Variable	VIF	1/VIF
TobinQ	2.35	0.424910
PB	2.10	0.475198
Lev	1.76	0.569510
SIZE	1.75	0.572978
StaffIntensity	1.23	0.816004
AverageAge	1.18	0.850930
FirmAge	1.13	0.885011
Digital	1.02	0.976297
Intangible	1.01	0.985248
FL	1.00	0.996761
Mean VIF	1.23	

According to the data shown in Table 1, it can be seen that the VIF value of each variable is less than 5, which meets the critical value requirement of multicollinearity. In addition, the Mean VIF value is 1.23, which suggests that the variables do not exhibit any issues with multicollinearity, so the subsequent regression results will be more reliable and accurate.

2.3 Model Specification

$Digital_{i,t}$ represents the primary dependent variable, reflecting the extent of digital transformation for the i-th company during the t-th period; CVs is the control variable group; $\sum Year$ is the time fixed effect; α_0 is the constant term; α_1 and $\sum\alpha$ are the regression coefficients of the corresponding variables; $\varepsilon_{i,t}$is the random error term.

$$ROE_{i,t} = \alpha_0 + \alpha_1 Digital_{i,t} + \sum\alpha CVs + \sum Year + \varepsilon_{i,t} \quad (1)$$

2.4 Theoretical analysis and research hypotheses

In essence, digital transformation precisely promotes the accumulation and application of digital resources for enterprises, which in turn brings new competitive advantages for enterprises to improve performance. Through digital talent recruitment and training, enterprises can improve the digital quality and innovation ability of the team, prompting enterprises to form a unique digital human resource advantage [8], creating favorable conditions for performance improvement.

Drawing upon the preceding analytical findings, the study formulates its first theoretical proposition:

Digital transformation of enterprises helps to improve financial performance (H1).

Existing research shows that SOEs often face the constraints of soft budget constraints and administrative governance in the process of digital transformation, which leads to resource mismatch and inefficient digital bureaucratization, thus weakening the enhancement effect of digital transformation on financial performance [8, 9]. In contrast, NSOEs, relying on market-based incentives, have more efficient resource allocation and flexible decision-making mechanisms, and are able to embed digital technologies into their business processes more effectively, significantly increasing their return on assets [9, 10]. There are empirical studies further suggesting that the effect of digital transformation on financial performance improvement in NSOEs is particularly significant in the eastern region, large enterprises and growth enterprises, while SOEs have had more limited results in digital transformation

compared to other types of enterprises due to rigid decision-making mechanisms and low resource allocation efficiency [5].

Drawing upon the preceding analytical findings, the study formulates its second theoretical proposition:

Due to differences in resource allocation efficiency and decision-making mechanisms caused by property rights attributes, the effect of digital transformation on financial performance is more significant in NSOEs than in SOEs (H2).

3 Analysis of empirical results

3.1 Descriptive statistics

According to Table 2 (all retained 4 decimal places),the average, minimum, and maximum values of return on net assets of the sample firms are 0.0517, -27.5949, and 1.1170, respectively, which represent significant differences in business performance among the firms. Some companies are actively adopting advanced digital technologies and tools such as artificial intelligence, big data analytics, blockchain, etc. And the application of these technologies improves operational efficiency and the quality of decision-making, which in turn enhances business performance. From the variable Digital, the average, minimum, and maximum values are 2.1995, 0.0000, and 5.0370, respectively, indicating that significant variations exist in the extent of digital transformation across businesses, with some enterprises yet to initiate any digital transformation efforts, and the overall advancement in this area remains relatively limited.

Table 2 Descriptive statistics results

Main variables	Representation	Mean	Std. dev.	Min	Max
Return on equity	ROE	0.0517	0.4498	-27.5949	1.1170
Digital transformation	Digital	2.1995	1.0787	0.0000	5.0370
Company Size	Size	22.2550	1.2332	18.4788	27.6377
Leverage Ratio	Lev	0.3939	0.1784	0.0091	0.9901
Average age of management	AverageAge	49.6910	3.2403	38.9206	62.8801
Years of establishment	FirmAge	2.9934	0.2961	1.3863	4.2195
Employee density	StaffIntensity	1.2602	1.1603	0.0215	65.8183
Price to Book Ratio	PB	3.8220	4.6277	0.1780	215.3516
Intangible assets ratio	Intangible	0.0417	0.0330	0.0000	0.6235
Tobin's Q Value	TobinQ	2.7071	2.0033	0.6873	44.3837
Financial Leverage	FL	1.1377	8.8951	-582.5801	547.8892

3.2 Benchmark regression

According to Table 3 (all values retained to four decimal places), regarding the influence of control factors, a negative relationship is observed between enterprise size and ROE (α=-0.0081, P=0.029), indicating that the phenomenon of “diseconomies of scale” is still prevalent in traditional manufacturing industries. Staff intensity shows a strong positive effect on performance (α = 0.0705, P<0.001), indicating that for labor-intensive companies undergoing digital transformation, the focus should be on enhancing human-machine collaboration. After controlling for the year and related control variables, the coefficient of enterprise digital transformation (Digital) is 0.00626, with a p-value of 0.052, which is close to 0.05, which can be illustrated to a certain extent to be significantly positive at the 5% level,

which suggests that with the gradual deepening of the enterprise's digital transformation, the enterprise's financial performance is also improved along with it, and the hypothesis H1 is valid.

Table 3 Regression results

ROE	Coefficient	Std	t	P>|t|	[95% conf. interval]
ROA	2.5429	0.0555	45.78	0.000	(2.4340, 2.6518)
TobinQ	0.1067	0.0026	40.57	0.000	(0.1015,0.1118)
Digital	0.0063	0.0032	1.94	0.052	(-0.0001, 0.01258)
PB	-0.0759	0.0011	-70.51	0.000	(-0.0780,-0.7380)
Size	-0.0081	0.0037	-2.19	0.029	(-0.0153, -0.0008)
LEV	0.4912	0.0255	19.26	0.000	(0.4412,0.5412)
AverageAge	0.0011	0.0012	0.97	0.334	(-0.0011,0.0034)
StaffIntensity	0.0705	0.0033	21.54	0.000	(0.0641,0.0770)
FirmAge	0.0437	0.0123	3.55	0.000	(0.0195,0.0679)
Intangible	0.1166	0.1051	1.11	0.267	(-0.0893,0.3225)
FL	-0.0006	0.0004	-1.64	0.101	(-0.0014, 0.0001)
_cons	-0.3594	0.0865	-4.16	0.000	(-0.5289, -0.1899)

3.3 Ownership heterogeneity test

First, the samples are divided into two groups according to their ownership structure: SOEs) and NSOEs, where the value '1' denotes SOEs. After eliminating missing values, there are 2115 observations. 0 represents NSOEs. After eliminating missing values, there are 7053 observations.

In this case, a high-dimensional fixed effects regression model (reghdfe) is applied for empirical analysis to examine how the core independent variable, Digital, influences the return on equity (ROE) for both SOEs and NSOEs.

$$ROE_{i,t} = \beta_0 + \beta_1 Digital_{i,t} + \beta_2 ROA_{i,t} + \cdots + \beta_k TobinQ_{i,t} + \gamma_t + \delta_j + \epsilon_{i,t} \quad (2)$$

Among them, the independent variables and the core dependent variables remain unchanged, and the year fixed effect （γ_t）and industry fixed effect（δ_j）are controlled to eliminate the interference of time trends and industry heterogeneity and avoid estimation bias caused by omitted variables. At the same time, by narrowing down the sample to include only SOEs and NSOEs and focusing on the unique characteristics of SOEs and NSOEs, the group explanatory power of the conclusion is enhanced. The results are shown in Table 4:

Table 4 Results of two-sample t-tests for group comparisons (assuming equal variances)

Group	Mean (SD)	Std.	95% CI	Mean Difference	t-value	p-value
0	2.23 (1.08)	0.0129	[2.21, 2.26]	0.1406	5.267	<0.001
1	2.09 (1.06)	0.0230	[2.05, 2.14]	–	–	–
Combined	2.20 (1.08)	0.0113	[2.18, 2.22]	–	–	–

The average value for group 0 and group 1 differ by a value of 0.1406. Since the P value is very small, less than the commonly used significance level (0.05), this paper can reject the

null hypothesis and believe that the difference in means between group 0 and group 1 (i.e., NSOEs vs. SOEs) is significant.

Table 5 Comparison of sub-sample regressions

Variable	Sample of SOEs(P-value)	Sample of NOSEs(P-value)	P-value for difference between groups
Digital	-0.0031 (0.077)	0.0086 (0.040)	$p < 0.01$
CVS	partly remarkable	mostly significant	

The results in Table 5 show that there is a difference in the impact of digital transformation on the financial performance of SOEs and NSOEs, supporting the hypothesis H2.

(1) In the category of privately-owned businesses, the digital sector plays a notable role in enhancing return on equity (ROE) ($\beta = 0.0086$, $p = 0.040$), suggesting that for every unit of improvement in their digital transformation, ROE increases by 0.86 percentage points on average;

(2) In the SOE group, Digital shows a non-significant negative association with ROE ($\beta = -0.0031$, $p = 0.077$), implying that digital transformation may not be effectively translated into financial performance improvement;

(3) "A test of difference in coefficients between groups using a seemingly unrelated regression (SUR) framework shows that the difference in DCG coefficients between SOEs and non-SOEs is significant at the 1% level ($\chi^2 = 4.62$, $p = 0.003$)."

4 Conclusion

This paper sample comprises manufacturing firms listed on the Shanghai and Shenzhen A-share mainboards, with data spanning the decade from 2013 through 2023, and adopts text mining and econometric methods to analyze the specific relationship between digital transformation and enterprise business performance. It is empirically tested that digital transformation can promote enterprises to enhance their own operational capabilities and then improve their operational performance, and the effect of ownership heterogeneity is significant. Based on this, the following policy recommendations are proposed:

1. Digitalization's nonlinear progression demands institutional adaptations wherein strategic policy incentives stimulate tech-business symbiosis, particularly through embedding intelligent systems in commercial processes and reengineering organizational hierarchies, consequently optimizing total factor productivity.

2. For SOEs, they should optimize their own transformation paths and establish a "digital transformation efficiency audit" system. A third-party agency should conduct an annual assessment of the return on investment (ROI) and technology penetration rate (such as AI application scenario coverage) of SOEs digital projects, and eliminate inefficient projects (such as duplicated ERP systems). At the same time, the digital transformation KPI should be included in the assessment system for SOEs executives, and quantitative indicators should be set (such as data center utilization rate >= 70%, automation equipment networking rate >= 90%), which should be directly linked to salary incentives.

3. Calibrated policy interventions tailored for privately held businesses must be intensified using digital maturity indices, realize the safe sharing of cross-enterprise data (such as real-time inventory in the supply chain and early warning data on equipment failures), and reduce the cost of data acquisition for SMEs.

For sure, this article also has many shortcomings. First, this article has indicator limitations: Digital is only measured by the frequency of annual report keywords, which may ignore the actual depth of technology application. In the future, patent data or IT investment ratio will be further supplemented to build a more comprehensive digital transformation

evaluation system. Second, there is a sample bias: unlisted companies and small and micro manufacturing companies are not included, and the universality of the conclusion is limited.

References

1. L. Wen. An empirical analysis of the effects and mechanisms of enterprise digital transformation on operating performance. National Circulation Economy **(21)**, 106-109 (2024)
2. China Academy of Information and Communications Technology. Research report on the development of China's digital economy (2024)
3. Z. Chen, S. Sun. Digital Transformation of Manufacturing Industry and New-Quality Productivity —Evidence from Listed Manufacturing Companies in Shanghai and Shenzhen Stock Exchanges
4. F. Wu, H. Hu, H. Lin & X. Ren. Enterprise Digital Transformation and Capital Market Performance: Empirical Evidence from Stock Liquidity (2021)
5. Y. Ren, Z. Tian, C. Liu, F. Ren & X. Guan. How Does Digital Transformation Affect the Performance of Private Energy Enterprises? Evidence from the Text Recognition of the Annual Reports of Listed Companies in China (2024)
6. C. Yuan, T. Xiao, C. Geng & Y. Sheng. Digital Transformation and Division of Labor between Enterprises: Vertical Specialization or Vertical Integration (2021)
7. G. Wang, M. Luo. The impact of digital transformation of business models on the financial performance of manufacturing enterprises: an analysis of the moderating effect based on technology embedding. Modern Marketing (First Half of the Month) **(01)**, 97-99 (2025)
8. X. Li, Z. Chen & Y. Chen. Digital Transformation of Enterprises and Financial Performance: Evidence from Digital Patent Data of Listed Companies (2024)
9. B. Yang, Z. Yang & Y. Guo. Employment structure effect in enterprises digital transformation: Empirical analysis based on text mining of annual reports of manufacturing industry listed enterprises (2023)
10. L. Kong. Digital Transformation and Development of Small and Medium-sized Enterprises in Digital Economy (2024)