New intelligent empowerment for digital transformation

Peng Yifeng [a, *], Gao Chen [b]

[a] School of Management and Economics, Chinese University of Hong Kong (Shenzhen), Shenzhen, Guangzhou 518100, China

[b] Rämistrasse 71, 8006 Zürich, Switzerland

Abstract

This study proposes an innovative evaluation method based on large language models (LLMs) specifically designed to measure the digital transformation (DT) process of enterprises. By analyzing the annual reports of 4407 companies listed on the New York Stock Exchange and Nasdaq from 2005 to 2022, a comprehensive set of DT indicators was constructed. The findings revealed that DT significantly improves a company's financial performance, however, different digital technologies exhibit varying effects on financial performance. Specifically, blockchain technology has a relatively limited positive impact on financial performance. In addition, this study further discovered that DT can promote the growth of financial performance by enhancing operational efficiency and reducing costs. This study provides a novel DT evaluation tool for the academic community, while also expanding the application scope of generative artificial intelligence technology in economic research.

Keywords: Digital Transformation, Corporate Innovation, Financial Performance, Fintech, Large Language Models

[*] Corresponding Author

E-mail addresses: yifengpeng@cuhk.edu.cn (Y. Peng), chengao@uzh.ch (C. Gao).

1. Introduction

The rapid development of information technology is continuously reshaping the economic landscape of modern society, with digital information and data-driven digital economy gradually becoming a significant driver of global economic growth. In this context, the deep integration and parallel development of the physical economy and digital economy have become the key theme in current economic evolution. In the face of the strong momentum of the digital economy, businesses and organizations are integrating digital technologies to revolutionize their business models and value creation channels, aiming to enhance operational efficiency, enhance customer experience, and promote innovative activities. This transformation process is widely recognized as "digital transformation" (DT). According to the World Bank's Digital Development Report (World Bank, 2024), more than 70% of global enterprises have completed or are actively progressing through the DT process. In 2023, the contribution ratio of the digital economy to GDP in 51 major economies across the world reached 46.1%, with a growth rate of 7.4%.

The necessity of corporate DT has become a common consensus among academia, with numerous empirical studies demonstrating that DT can significantly enhance the performance and business continuity of companies (George & Schillebeeckx, 2022; Guandalini, 2022), while also having a positive effect on promoting corporate innovation (Appio et al., 2021; Drechsler et al., 2020; S. Li et al., 2023). Nonetheless, a survey study conducted among corporate executives revealed that only approximately 30% of digitally transformed companies were rated as "successful" or "having significant impact" (Jacques Bughin et al., 2019). The vast discrepancy between the promising prospects of DT and the practical challenges it faces stems from multiple factors. Among them, corporate executives may misinterpret the initial efforts to drive change as a breakthrough for market dominance, leading to erroneous assessments of the situation; excessive investment in infrastructure may cause most companies to be bogged down by heavy burdens of funds and resources; at the same time, a lack of clear evaluation standards for the success of new technologies may leave companies feeling lost in terms of strategic direction (Davenport & Westerman, 2018; Oludapo et al., 2024).

This conceptual difference between academia and practical business reveals the core difficulty of the current DT, which is the lack of a unified standard for evaluating corporate DT. Restricted by the background and industry differences of corporations, previous research results often differ in comparability. Current popular standards for evaluating corporate DT include word frequency statistics, questionnaire surveys, and text analysis using built-in dictionaries (Larysa Yakymova et al., 2022; Xie & Wang, 2023). However, current text analysis methods in accurately measuring corporate transformation progress pose limitations. In particular, when constructing DT-related corpora, this method primarily relies on historical text data, which leads to significant lag in the rapidly evolving IT field. Additionally, ignoring contextual text analysis may misrepresent some keywords as the hallmark of DT. This is particularly evident in corporate financial reports, where companies may frequently mention advanced technologies, but this does not necessarily indicate that these technologies have been applied and empowered in production or management processes.

To address the shortcomings of the current digital standards, this study adopts a data processing strategy based on large language models (LLMs). This approach replaces the traditional single text analysis technique and fully harnesses the potential of contextual information in text, allowing for in-depth analysis of corporate annual reports. The study is based on corporate annual reports from companies listed on the New York Stock Exchange and Nasdaq Stock Exchange from 2005 to 2022, encompassing 4,407 listed companies. The construction of the DT standard system can be divided into the following steps:

Step one, using web crawler programs and manual collection methods to acquire annual reports from listed companies, with a focus on the content of "Management's Discussion and analysis" and "Principal Risk/Risk Management/Risk Factor" in the annual reports' text. Those content can be considered as strong-correlated financial texts with the corporate DT (Pramanik et al., 2019).

Step two, all relevant texts are split into sentences and formed into a sentence library for prediction.

Step three, simultaneously extract relevant texts and texts containing keywords, forming a sentence library for annotation. The annotated sentence library is used to evaluate whether the company has undergone DT.

Step four, based on open-source large language models, a supervised machine learning method is adopted to train a sentence classification model.

Step five, use the trained large language model to predict each sentence in the prediction sentence library, and judge whether listed companies use and use which type of digital technology. Finally, a new set of indicators for enterprise DT is constructed.

In order to confirm the applicability of the constructed indicators, this study conducted multi-level comparative analyses successively. Based on these indicators, in-depth comparative research was conducted on patent data, regional data, and international literature, etc. The results showed that the indicators constructed in this study highly aligned with reality. Compared with the dictionary method, the indicators constructed in this study demonstrated superior performance in terms of content completeness and semantic authenticity.

Based on the construction of the enterprise DT indicator system using large model methods, this study explores the relationship between enterprise DT and enterprise financial performance through empirical analysis, revealing three important findings. Firstly, overall, DT has a significant positive impact on improving financial performance (i.e., asset return rate (ROA) and equity return rate (ROE)). However, not all applications of digital technologies can significantly enhance financial performance: specifically, applications of big data(BD), artificial intelligence(AI), mobile internet(MI), cloud computing(CC), and the Internet of Things(IoT) all significantly improve ROA and ROE, whereas applications of blockchain technology have not been able to demonstrate a significant positive impact on ROA and ROE. Secondly, the effects of DT on financial performance vary among different financial performance enterprises. For enterprises with poor financial performance, DT can effectively enhance their ROA and ROE. However, for enterprises with good financial performance or especially outstanding performance, the impact of DT on ROA and ROE is not significant. Further, the primary ways in which businesses enhance their financial performance through DT include improving operational efficiency and reducing operational costs, while empirical support for revenue growth remains absent.

The primary contributions of this study can be summarized into three aspects. Firstly, this study provides an innovative measurement method for assessing corporate DT, thereby establishing a solid empirical foundation for subsequent related topics. Previously, studies on the impact of DT on corporate behavior and performance often began from the purchase of digital devices and services by enterprises (Guo et al., 2023; Pereira et al., 2023). However, due to the differences in measurement objects and methods, the results of these studies could not be effectively compared, and conclusions drawn by different literatures might be mutually conflicting (Mielli & Bulanda, 2019). This study, based on large language models, provides a comprehensive, accurate, and replicable set of indicators for measuring DT, characterized by clear objectives, complete indicators, high accuracy, and replicability. This strategy addresses key issues in current research, thereby deepening the study of corporate DT and providing empirical evidence for the broad literature on digital economy.

Secondly, this study elucidates the distinct impacts different digital technologies have on corporate financial performance, and identifies various impact pathways, providing a novel perspective for the existing literature on DT. In previous research, the assessment of DT often lumps the application of all digital technologies into the symbolic representation of corporate DT. This approach carries certain drawbacks. In actuality, different digital technologies might yield different transformation effects, and enterprises may only adopt one type or a few types of these technologies in practice. Based on Vial (2019) and CAICT (2024) comprehensive literature review, this study categorizes digital technologies into six major types: artificial intelligence, big data, the internet of things, cloud computing, mobile internet, and blockchain. The results reveal that compared to the other five types of digital technologies, blockchain technology has a limited positive impact on corporate financial performance. Simultaneously, other five types of technologies significantly enhance corporate financial performance. Furthermore, this study validates the effectiveness of efficiency enhancement channels and cost reduction channels, while the hypothesis of revenue enhancement channels is not validated. This finding aligns with the existing academic literature results. By categorizing digital technologies more finely and verifying the impact pathways, this study provides a new perspective and contribution to the literature on corporate DT.

Thirdly, this study expands the application scope of generative artificial intelligence technology in the field of economics. With the rapid advancement of computer science technology, numerous researchers have begun to utilize LLMs to conduct in-depth discussions on microeconomic issues. Particularly, following the implementation of the BERT model by Dogu (2019) to perform sentiment analysis and market trend prediction on financial text data, a series of optimized models named FinBERT have emerged (A. H. Huang et al., 2023; Yang et al., 2020; Yifeng, 2024; Y. Zhang & Zhang, 2023). Zhou et al. (2023) infused enhanced search components into the large language model, enabling the supplementation of contextual information in financial text and thus facilitating fine-grained instruction tuning. On this basis, Tong et al. (2024) further expanded the research on financial text and financial data to multiple modalities, significantly enhancing the performance of the model in quantitative strategy generation. Notably, in the field of corporate DT research, this study has for the first time attempted to utilize LLMs for quantitative analysis. This study successfully opens a new exploration path for the application of large language models in the field of economics research.

The contents of the subsequent sections will unfold according to the established structure. In the second section, a comprehensive evaluation of indicators for enterprise DT will be conducted. In the third section, the construction of these indicators will be presented, focusing on how they can be built using machine learning technologies and large language models. The validity of these indicators will be validated from various perspectives. In the fourth section, a deep analysis of enterprise DT performance will reveal a series of new findings. The conclusion and outlook of the research results will be summarized and discussed at the end.

2. Literature Review on DT

Existing literature performed three main methods when assessing the degree of corporate DT. Firstly, some studies start with publicly available data, calculating a series of indicators through statistical computation to evaluate the degree of corporate DT. For instance, the assessment of investments in digital technologies related to software or hardware and their proportion to total assets (Ghobakhloo, 2020), the emphasis on investments in information services (Noy & Zhang, 2023), or the scrutiny of data analysis software services (Pappas et al., 2018), as well as the observation of information technology density (Nwankpa & Datta, 2017). However, the two main shortcomings of this approach should not be overlooked. Firstly, its applicability is relatively narrow, only capable of measuring the non-human cost of a specific digital technology. For example, if a company employs engineers dedicated to DT, the wages paid to these engineers should be viewed as part of the company's DT investment. However, such investments do not reflect in the costs of digital hardware or software. Secondly, the measurement method is relatively rough. By simply counting digital technology hardware or software, it is not possible to differentiate between the specific applications of different types of digital technologies.

The second method focuses on exploring how specific events or policies promote the DT of businesses. Satish et al. (2019) proposed a social framework for the interaction between government incentives and digital innovation, suggesting that differences in innovation performance can be explained through differences in local government policies on innovation (J. Li et al., 2018). Moreover, the interactive relationship between the government and businesses is of crucial importance for understanding the endogeneity issues in the process of DT. The limitations of the event study method also exist. Firstly, different companies exhibit varying responses and adaptive capabilities to the same policy, which may lead to varying degrees of their exposure to the policy. Secondly, businesses within the pilot region may simultaneously be affected by other policies in the region. Although the interference effects of other policies in the sample period can be eliminated through parallel trend tests, it is difficult to eliminate the impact of other policies implemented during the same period as the pilot policy. In addition, the concurrent launch of multiple related policies, coupled with the lag characteristics of policy effects, makes it more challenging to differentiate the specific effects of different policies.

The second method focuses on examining how specific events or policies promote the DT of businesses. Satish et al. (2019) proposed a social framework that explores the interactive effects of government incentives and digital innovation, suggesting that differences in innovation performance can be explained by different approaches adopted by local governments in innovation policies (J. Li et al., 2018). Additionally, the interactive relationship between the government and businesses holds significant implications for understanding the endogeneity issues related to corporate behavior and performance during the transformation period (Wang et al., 2023). However, this approach also has its limitations. Firstly, different businesses may exhibit varying responses and adaptability to the same policy, which could lead to inconsistencies in their level of impact. Secondly, businesses within the pilot region may be simultaneously influenced by other policies in the region, making it challenging to eliminate the interference effects of other policies during the sample period. Furthermore, the simultaneous release of multiple related policies and the characteristic of policy

effects' latency make it more difficult to differentiate the specific effects of different policies.
The third approach is the widely adopted text analysis technology at present. Once a specific digital technology is adopted by an enterprise, the frequency or proportion of its mention in their public reports will correspondingly increase. By analyzing publicly available information from listed companies, leveraging text analysis technology, we can quantify the occurrences of keywords or key sentences and their trends, thereby establishing a metrics system to assess the level of DT within the company (Guo et al., 2023; Tavana et al., 2022; Zhai et al., 2022). Although keyword-based dictionaries are widely used, there are inevitable delays in updating and bias introduced by subjective human choices during keyword construction. It is worth noting that the dictionaries may mistakenly include unrelated information regarding enterprise DT practices into consideration. In interpreting company financial reports, mere occurrences of keywords cannot be taken as conclusive evidence that the company has carried out corresponding DT in the absence of context. This is mainly due to several reasons: firstly, the company may simply be planning future DT, and has yet to execute it; secondly, keywords may appear in a context of negation; lastly, the company may merely be describing the development background of the industry, rather than its own DT actions. In the review of DT, it is frequently emphasized that the application of digital technology by enterprises without being recognized belongs to the first type of error, while the mention of keywords in the text without actually using digital technology belongs to the second type of error (Agostino & Costantini, 2021; İnel, 2019; Tavana et al., 2022). This study aims to use a method based on LLMs to conduct in-depth research on the above-mentioned problems and propose corresponding optimization solutions.

3. Measurement Based on LLMs

3.1 Llama3 Model Introduction

The application of dictionary-based text analysis techniques, although convenient in operation, is inherently limited in the extraction of contextual information related to keywords. It cannot effectively delve into the semantic meaning at the sentence level. In recent years, with the rapid development of artificial intelligence technology, pre-training model techniques have become particularly crucial in the field of natural language processing (NLP). Specifically, pre-training refers to the initial training of a basic model using a large-scale dataset, followed by fine-tuning on a particular task to improve the accuracy of that task. Pre-training techniques can be categorized into early static pre-training techniques and the more advanced dynamic pre-training techniques. The primary difference between the two lies in the contextual sensitivity of word representations. Dynamic pre-training techniques, exemplified by the large language models GPT and BERT (Devlin et al., 2019; Radford et al., 2018), innovatively propose a method for text representation based on context, effectively addressing the issue of polysemy. As technology evolves further, Sun et al. (2020) enhance the model's semantic representation ability by introducing external knowledge through masking operations on words (i.e., knowledge units). By May 2024, META's open-sourced LLaMA3 model surpassed the latest technical level of advanced models in various industry benchmark tests, and its performance was superior across numerous standard test benchmark standards (W. Huang et al., 2024; META, 2024). Consequently, this study selected the LLaMA3-8B model to perform text classification tasks.

3.2 Using the LLaMA3 to measure firm DT

3.2.1 Determine the object of text analysis

DT represents a comprehensive revolution activity across multiple dimensions, including organizational structure, internal management systems, and business processes, which may not readily manifest its impact on financial performance indicators (Davenport & Westerman, 2018). Despite this, Asikpo (2024) found through text analysis of annual reports from listed companies that these entities commonly exhibit a strong motivation to disclose relevant information in their annual reports, in order to garner recognition and preference from the capital market. Based on this observation, this study employs annual reports from listed companies as the text analysis data to analyze the DT process of these companies.

This study employs web scraping technology and combines manual review, extracting relevant textual data from the annual reports of 4407 publicly listed companies on the New York Stock Exchange and the NASDAQ Stock Exchange from 2005 to 2022. In the "Management Discussion and Analysis" (MD&A) section of these annual reports, which typically contain in-depth analysis of the company's current operations, detailed explanations of its future strategies, and clear disclosures of the risks it faces, we found that some companies tended to announce potential risks they might face in "Primary Risks/Risk Management/Risk Factors" sections of the annual reports. Among these risks, there may be information related to the DT of the company. Finally, this study compiled and organized 48,547 annual report texts from 4407 companies from 2005 to 2022.

3.2.2 Building a database of predicted and marked sentences

For the collected texts, it is necessary to use a sentence segmentation algorithm to divide them into sentences and paragraphs. The resulting sentence pool should be used to predict the sentences. Since most of the texts in annual reports are unrelated to DT, if sentences are randomly selected for reading, most labels obtained will be unrelated to digital technology. To improve the efficiency of manual reading and prevent contextual interference, it is necessary to extract keywords from the annual reports that represent different natures, and combine them with randomly selected sentences to form the sentence pool for labeling. Therefore, it is necessary to identify the definition of digital technology clearly and build a digital technology dictionary.

Digital technology, as a broad and continuously evolving field, encompasses various aspects from basic computational techniques to the various levels of complex networks and communication systems. This study draws from policy documents, research reports, and existing literature (Ghobakhloo, 2020; İnel, 2019; Wang et al., 2023), and constructs a word list containing 487 terms related to digital technology through manual deep reading and continuous supplementation. To enhance the model's accuracy in handling sentences without keywords, we randomly extracted portions from annual reports and performed sentence-level segmentation. Considering that the number of listed companies is increasing annually, if we directly annotate these sentences, the sentences will be highly biased toward recent years. To address this issue of imbalanced year distribution, we further grouped these sentences by year, ensuring an equal number of sentences per year. Then, from these sentences evenly distributed across the years, we perform an unbiased random sampling without replacement, forming a sentence collection for annotation. Finally, the annotated sentence collection consists of 46,754 sentences.

3.2.3 Manual Labelling

The primary objective of manual labelling is to construct training sets, test sets, and validation sets to lay the foundation for subsequent machine learning tasks. During the annotation process, it is necessary to identify the type of digital technology adopted by the enterprise, and assess whether the enterprise has undergone DT.

In this study, we invited 32 academic researchers with backgrounds in financial technology, and they were divided into groups of two individuals each. These groups would rotate regularly. In the actual annotation process, each sentence in the annotated corpus would be annotated by two researchers. If the two researchers agreed on the annotation results, the sentence would be labeled accordingly; however, for sentences with disagreements in labeling, we would determine the final label through group discussion. For sentences that were difficult to label, we would not include them in the training set. Finally, in addition to sentences that were difficult to label, the rest of the annotated sentences in the corpus were classified into eight labels. These labels include six new digital technologies, non-new digital technologies, and non-digital technologies.

3.2.4 Training with Supervised Learning

The novelty of this study lies in the utilization of LLMs to replace human evaluation of whether the DT keywords present in text represent the actual implementation of digital technology by the enterprise, thus addressing the issue of texts that contain digital technology keywords but do not implement digital technology in reality. In data partitioning, this study divides all annotated sentences into training set, test set, and validation set, in a ratio of 8:1:1. During model training, we utilized the tokenizer function built into LLaMA3 to quickly convert sentences into the formats necessary for model training, and to save training time, all data were trained using Low-Rank Adaptation (LoRA), a rapid fine-tuning method (Hu et al., 2021).

The core objective of machine learning is to identify whether a text reflects a specific digital technology, as well as which type of digital technology is being specifically demonstrated. When evaluating the performance of classification models, commonly used measures include Precision, Recall, and Accuracy. Due to the possibility of imbalanced label counts in the training dataset, F-Score is often used to assess the overall classification performance of the model. In this study scenario, annual reports may contain multiple instances of sentences indicating the use of digital technologies by the company. Therefore, the ability to accurately classify becomes particularly important in this context. In light of this consideration, this study further calculates the F.7-Score, thereby assigning a higher weight to Precision during the evaluation process. Table 1 presents the performance of each model on the same training set.

Table 1 Comparison of classification performance

	Accuracy	Precision	Recall	F1 Score	F0.7 Score
GaussianNB	70.13%	49.63%	52.63%	0.5109	0.5076
SVM	80.34%	65.75%	56.29%	0.6065	0.6170

Voting	81.13%	66.59%	58.44%	0.6225	0.6315
NN	79.74%	62.96%	61.89%	0.6242	0.6254
KNN	79.28%	64.12%	54.72%	0.5905	0.6009
BERT	89.32%	76.84%	73.56%	0.7517	0.7553
LLaMA3	92.61%	82.77%	76.35%	0.7943	0.8014

Note: The BERT algorithm used for comparison is trained on the PaddleHub framework, while the other machine learning methods are trained on the sklearn library.

3.2.5 Constructing Digital Transformation Indicators by LLaMA3

Based on the LLaMA3 language model, this paper predicts texts in the corpus of sentences to be predicted from 2005 to 2022, and judges whether a company has used digital technology and which type of digital technology it has employed. It also constructs a dummy variable for corporate DT, indicating that if a company has employed any of the technologies of AI, BT, CC, BC, IoT, MI in any year, the indicator assignment is 1, otherwise it is 0.

3.3 Validation Test

3.3.1 Classification Performance

Compared to the widely adopted dictionary method in the current literature, the LLaMA3 model demonstrated superior performance in various evaluation metrics. To illustrate this, this study selected Chawla (2021) and Kraus (2022) as comparative baselines. These two studies, respectively based on 93 and 161 keywords, conducted classification performance analysis using dictionary methods. The results indicated that although the latter study employed a larger number of keywords, this did not significantly improve the classification performance. The differences were more apparent in the granularity of domain-level segmentation. This finding emphasizes the advantages of adopting large-model methods in accurately assessing whether a text reflects the DT of a company, i.e., the significant effectiveness in enhancing the sincerity of the expression.

Compared to dictionary-based methods, strategies based on LLMs exhibit an optimization of nearly 30% in handling false positives, that is, issues mentioned but not utilized in reality, which reflected in the change in Precision values. Additionally, the strategy also achieves an optimization of 6-7% in handling false negatives, that is, issues where the implied use of digital technology cannot be captured by the model or keywords, which reflected in changes in Recall values. Those data clearly demonstrate that the methods employed in this study outperform dictionary-based approaches in correcting errors in both false positives and false negatives.

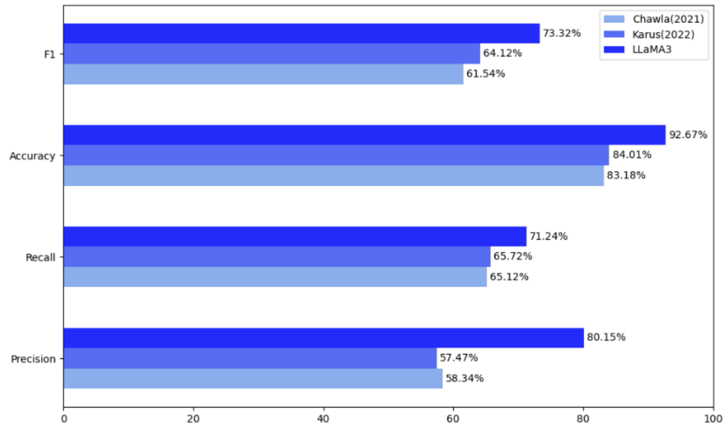

Fig. 1. Comparison between LLaMA3 and Dictionary approach

3.3.2 Time Trend

By calculating the cumulative mean of the dummy variable representing whether or not a listed company adopted a specific digital technology in a given year (as shown in Figure 2), we can gauge the popularity of various digital technologies across the years. From time series analysis, it is evident that as time progresses, the popularity of different digital technologies has significantly increased, particularly between 2010 and 2019. This observation aligns with the objective reality of technological advancement. From the perspective of the trend in technology adoption, overall, the adoption rates of IoT and AI technology are the highest, at approximately 60%, followed by big data and mobile internet technology, with an adoption rate of approximately 40%. Cloud computing technology ranks third, with an adoption rate of 20%, and blockchain technology has the lowest adoption rate, at 7%. This finding aligns with the conclusions of Xie (2023) and others, that the popularity of digital technology is closely linked to the time of its technical implementation.

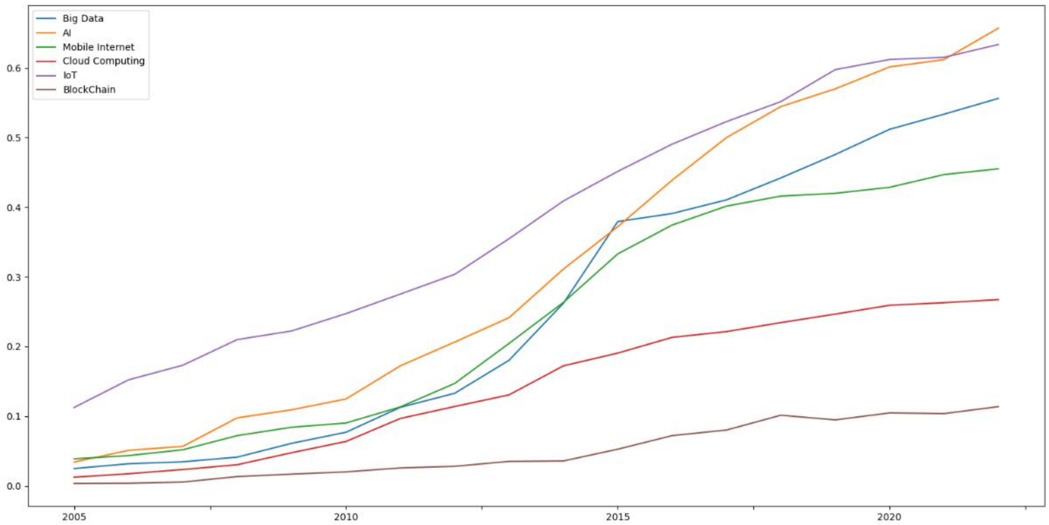

Fig. 2 Popularity of digital technology

3.3.3 Industry perspective

Based on the industry classification standards set by the North American Industry Classification System (NAICS), this study quantifies the application of digital technologies in various industries (Table 2). The analysis reveals that information transmission, software, and information technology services, finance, and science and technology services are leading in terms of DT. Conversely, agriculture, forestry, fishing, mining, and utilities industries appear to lag behind in the digital process. According to The World Bank's 2023 Digital Economy report(The World Bank, 2023), the digital economy penetration rate in developed countries in the agriculture, industrial, and service sectors, respectively, are 13.3%, 33.0%, and 46.7%. This data corroborates the accuracy of the measurement methods employed in this study.

Table 2 Popularity of digital technology in various industries

Industry name	AI	BD	CC	IoT	BC	MI
Information	71%	72%	61%	72%	16%	71%
Finance and Insurance	54%	59%	30%	24%	37%	41%
Professional, Scientific, and Technical Services	52%	48%	23%	65%	2%	24%
Arts, Entertainment, and Recreation	39%	42%	24%	31%	3%	60%
Management of Companies and Enterprises	36%	57%	37%	16%	26%	34%
Health Care and Social Assistance	35%	35%	20%	36%	2%	37%
Manufacturing	35%	23%	10%	54%	2%	21%
Educational Services	34%	25%	15%	37%	<1%	23%
Accommodation and Food Services	36%	40%	7%	27%	2%	38%
Retail Trade	34%	35%	10%	24%	3%	34%
Transportation and Warehousing	32%	33%	8%	41%	4%	18%
Wholesale Trade	31%	30%	9%	29%	2%	8%
Construction	28%	21%	10%	51%	5%	17%
Real Estate and Rental and Leasing	21%	27%	8%	20%	<1%	14%
Utilities	19%	17%	5%	33%	<1%	7%
Administrative and Support and Waste Management and Remediation Services	18%	21%	7%	16%	1%	9%
Public Administration	17%	18%	6%	8%	3%	18%
Mining, Quarrying, and Oil and Gas Extraction	16%	12%	3%	30%	8%	5%
Other Services (except Public Administration)	17%	16%	8%	6%	<1%	10%
Agriculture, Forestry, Fishing and Hunting	14%	14%	2%	27%	<1%	7%
Mean	33%	31%	15%	32%	6%	24%

Note: In the Mining, Quarrying, and Oil and Gas Extraction industries, the apparent high rate of adoption of blockchain technology is actually due to the overlap between the "mining" virtual currency behavior in the blockchain technology and the core keywords related to the mining industry. This does not reflect the widespread adoption of this technology by related businesses.

3.3.4 Patents Application perspective

In the process of enterprise DT, it is common to seek protection for relevant patents to emphasize the technological advantages of the company. If a company has already applied for a patent related to a certain type of digital technology, we can infer that it has implemented corresponding digital technology. However, reverse inference is not always possible. Therefore, we can compare the prediction results of various models with real companies that have applied for digital technology patents. The model whose prediction results are most consistent with the actual patent application situation can be deemed the optimal model.

The matching work between patent databases and corporate databases will be conducted following the following procedure: Firstly, the patent search platforms such as WIPO, FPO, and USPTO will be utilized to identify the patent application records of listed companies in the field of digital technology. Secondly, based on the World Intellectual Property Report 2022: The Direction of Innovation released by the World Intellectual Property Organization (WIPO, 2022), the patent classification numbers corresponding to the six categories of digital technology will be determined. Finally, by matching the patent's "standard applicant" field, these patents can be correctly assigned to the corresponding listed companies.

Compared with the dictionary method (Chawla & Goyal, 2021; Kraus et al., 2022), the LLaMA3 model demonstrates the highest level of consistency in identifying companies that have applied for patents in the field of digital technology (Figure 3). This finding indicates that the LLaMA3 model adopted in this study exhibits the highest level of accuracy when evaluating the level of corporate DT from a perspective of patent data.

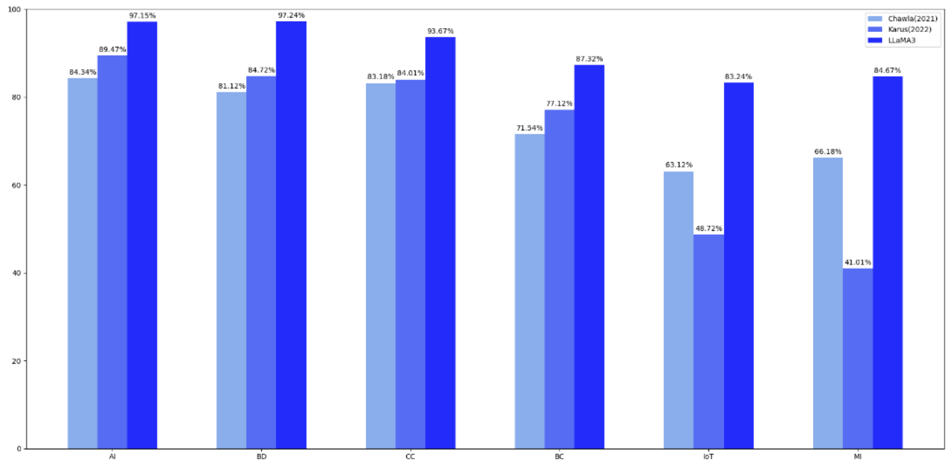

Fig. 3. The overlap between the recognition results and the company's patent applications

Note: The keywords extracted by Kraus(2022) did not specifically subdivide the direction of IoT and MI technology. We manually screened relevant vocabulary in professional dictionaries and manually classified them according to the aforementioned technology directions.

3.3.5 International Comparison

According to data from the National Bureau of Economic Research (NBER, 2023; 2020), there is a positive correlation between firm size and the adoption of AI technology. Specifically, among firms with the largest size, the adoption of AI technology accounts for 78%, while among firms with less than 50 employees, the adoption rate is less than 15%. By selecting data from 2020 and grouping firms based on employee size, Figure 3 reveals the correlation between enterprise size and AI usage rate. The artificial intelligence indicators constructed in this study exhibit high consistency with the statistical data in characterizing enterprise size.

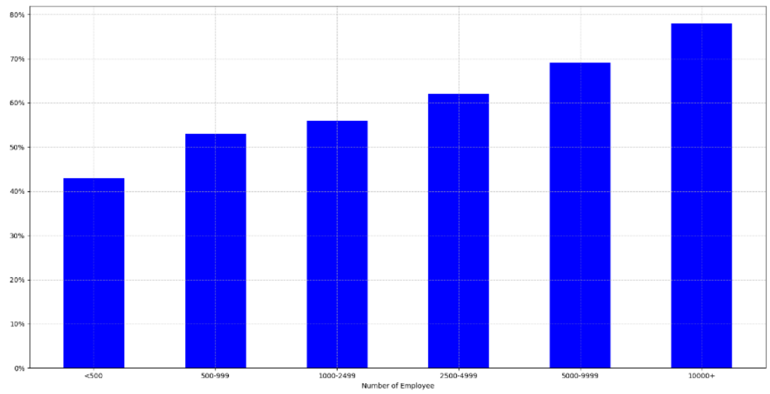

Fig. 4. The proportion of AI usage by listed companies in 2020

4. Digital Transformation and Financial Performance

To further validate the rationality of the enterprise DT indicators constructed in this study and provide a new perspective on the differences in the success of enterprise DT, this chapter will focus on analyzing the impact of enterprise DT on corporate financial performance. Choosing corporate financial performance as the dependent variable in this study is based on its objective nature, ease of measurement, and high comparability.

4.1 Regression models and variable definitions

Numerous academic studies have explored the impact of enterprise DT on corporate financial performance from various perspectives. However, in existing research, there is no exploration of the differences in how different digital technologies affect corporate financial performance. To understand the success or failure of DT, it is not only necessary to consider the overall DT, but also to differentiate how different digital technologies affect financial returns. This allows for a thorough examination of the pathways and mechanisms through which enterprise DT influences corporate financial performance.

In order to delve deeper into the impact of enterprise DT on corporate financial performance, this paper proposes the following benchmark regression model:

$$FP_{i,t} = a + b \cdot DT_{i,t} + \sum_{n} F_n \cdot Controls_{i,t} + \gamma_t + \nu_i + \varepsilon_{i,t} \tag{1}$$

In formula (1), the dependent variable $FP_{i,t}$ represents the financial performance of Firm i in year t, which is measured through ROA (total asset returns) and ROE (net asset returns). The key explanatory variable $DT_{i,t}$ is a group of dummy variables used to measure the degree of DT within the firm. This includes whether the firm has undergone DT, as well as whether it has adopted at least one new digital technology. $Controls_{i,t}$ represents a series of control variables. Referencing the existing literature, this study includes variables such as firm size, firm age, revenue growth (measured by the ratio of market capitalization to book value), the proportion of shares held by the largest shareholder, as well as cash flow, in the regression equation to control for these variables. In this model, γ_t represents time-fixed effects, and ν_i represents firm fixed effects. The random disturbance term is characterized by $\varepsilon_{i,t}$. This paper not only analyzes the DT indicators of firms, but also utilizes other related variables provided by Bloomberg and Wind databases.

Considering the significant impact on business operations caused by the global financial crisis and the COVID-19 pandemic, the regression model in this study selects the data from 2010 to 2019 for analysis. During the data processing process, we excluded financial institutions and samples with missing key variables, and we also performed 1% trimming on continuous variables. After these steps, we obtained 28041 observations. As shown in Table 3 and Table 4, the descriptive statistics indicate that 70.46% of the sampled companies adopted at least one new digital technology for DT, with the highest application rate for IoT and AI. In terms of financial performance, the average ROA for listed companies was 6%, and the average growth rate of revenue was over 14%.

Table 3 Explained Variable

Variable	Mean	Std	MAX	MIN	Observation
ROA	8.387	10.284	31.417	-20.892	28041
ROE	13.113	17.372	47.110	-51.024	28041

Table 4 Explanatory Variable

Variable	Mean	Std	MAX	MIN	Observation
Digital Transformation	0.7046	0.4271	1	0	28041
Artificial Intelligence*	0.4301	0.4667	1	0	28041
Big Data*	0.3196	0.3912	1	0	28041
Cloud Computing*	0.1733	0.3042	1	0	28041
Internet to Thing*	0.4712	0.5071	1	0	28041
Blockchain*	0.0708	0.1801	1	0	28041
Mobile Internet*	0.2948	0.3814	1	0	28041
Firm Age	3.204	0.437	5.370	0	28041
Firm Asset	6.827	0.613	8.072	5.248	28041
Firm Revenue Growth	17.39	41.03	170.02	-89.63	28041
Market to Book	0.6314	0.266	1.093	0.0873	28041
Cashflow/Asset	0.0383	0.0594	0.2291	-0.1730	28041

Note: * Indicates whether a certain technology is being used，firm age and asset is calculated by adding 1 to the number and taking the logarithm, firm asset counted by 1,000$.

4.2 Benchmark regression

The results of the benchmark regression are detailed in Table 4, with the primary explanatory variable defined as the dummy variable for corporate digital transformation (referred to as DT). Whether the dependent variable is ROA or ROE, the primary explanatory variable consistently demonstrates a significant positive relationship at the 1% level. This finding validates the significant improvement in corporate financial performance through the adoption of a new indicator constructed based on the LLaMA3 model. However, it is important to note that the benchmark regression analysis may involve reverse causality issues, wherein companies with better financial performance and more robust cash flows are more likely to adopt digital technologies. To mitigate potential reverse causality issues, dependent variables of the ROA^{-1} and ROE^{-1} rows in Table 5 were all lagged by one period. The results revealed a significant positive relationship between the corporate DT and the primary explanatory variable, which supports the initial finding that corporate DT, in general, promotes an overall improvement in corporate financial performance. Given that most companies in the information transmission, software, and information technology services industries belong to the digital industry, their use of digital technology differs from that of companies in other industries. Therefore, we excluded these companies from the regression analysis in Table 6. The results still aligned with the benchmark regression analysis, further strengthening the robustness of the aforementioned conclusions.

Table 5 Benchmark regression results of enterprise digital transformation and financial performance

Variables	DT	Controls	Annual FE	Firm FE	Observation	R^2
ROA	0.4372^{***}	YES	YES	YES	28041	0.2213
	(0.1027)					
ROE	0.7235^{***}	YES	YES	YES	28041	0.1663
	(0.2490)					
ROA^{-1}	0.1576^{*}	YES	YES	YES	24703	0.1169
	(0.0887)					
ROE^{-1}	0.5081^{**}	YES	YES	YES	24703	0.0787
	(0.2788)					

Note: The standard error in parentheses is defined as the clustering standard error at the enterprise level, * p<0.10 ** p<0.05 *** p<0.01.

Table 6 The impact of enterprise's use of DT on ROA and ROE (excluding information transmission, software, and information technology services)

	ROA	ROE
Digital Transformation	0.3432^{***}	0.7031^{***}
	(0.0749)	(0.2331)
Firm Age	-7.983^{***}	-11.502^{***}
	(2.771)	(3.374)
Firm Asset	1.580^{***}	3.232^{***}
	(0.2031)	(0.249)
Firm Revenue Growth	0.0776^{*}	0.1012^{**}
	(0.03703)	(0.03033)
Market to Book	-6.124^{***}	-15.647^{***}
	(0.341)	(1.981)
Cashflow/Asset	9.343^{***}	17.07^{***}
	(0.4112)	(2.9571)
Annual FE	YES	YES
Firm FE	YES	YES
R^2	0.1973	0.1443

4.3 Endogeneity Issues

In the context of regression models, apart from reverse causality, the problem of missing variables may also arise. To more effectively address endogeneity issues, refer to the approach of Tania et al. (2024), viewing the correlation between enterprises and top-tier schools in the field of information technology as a tool variable (IV) for corporate DT. For companies undergoing DT, the presence of insufficient skills among existing employees may pose a challenge. The recruitment of employees from top-tier schools can be considered a crucial factor.

The higher the proportion of employees recruited from these schools, the greater the likelihood of a company undergoing DT.

This study selected key terms such as Artificial Intelligence, Data Science, Mobile Computing, Cloud Computing, etc. from the CSRankings (2023) and narrowed down to a group of 30 universities in the field of Information Technology with advantages in North America. This group was then referred to as the "Leading University of IT" group. Generally speaking, the greater the average distance between the office locations of these "Leading University of IT", the more difficult it is for companies to attract talented individuals, thereby increasing the difficulty level of DT. Conversely, the greater the number of companies in a specific region, the more likely it is that each individual company is affected by local policies, thereby reducing the possibility of DT. Therefore, the geographical distance of the pilot schools and the number of companies in the region are negatively correlated with the likelihood of companies undergoing DT. This result conforms to the assumption of instrumental variable correlation. The definition of instrumental variable used in this paper is as follows:

$$IV_{i,t} = \frac{\sum_{j=1}^{30} DisFirm_i}{30} \times \frac{Count_{i,t}}{\rho} \tag{2}$$

In formula (2), i represent for firm i, t for the year, $DisFirm_i$ indicate the straight-line distance from the registered official office address of listed company i to the main campus of school j(Most University has multiple campuses, choose the one with the highest number of students as the main campus, distance calculated in 1000km). $Count_{i,t}$ means total number of listed firms in the city where company i is registered in year t. ρ represent for a constant number for calculation.

Table 7 presents the regression results of the instrumental variables method. Here, the coefficient of the instrument variable is negative and F-statistics exceed 10, which aligns with the expectation of correlation. It can be concluded that the strong comprehensive group itself does not directly impact corporate DT. Therefore, this instrument variable satisfies the exogeneity assumption. Simultaneously, the regression results of ROA and ROE indicate that after utilizing the explanatory variable, the coefficient of corporate DT is significantly positive at the 5% level. This suggests that after eliminating potential endogeneity issues using the instrumental variables method, the conclusions of this study are robust.

Table 7 Regression results of instrumental variable method

	Digital Transformation	ROA	ROE
Instrumental Variable	-0.0892**		
	(0.0356)		
Digital Transformation		11.354***	24.732**
		(5.024)	(15.612)
Wald F-statistic	23.015		
Wald rk F-statistic	14.422		
R^2		-0.172	-0.288

Note: Wald F-statistic refers to Cragg-Donald Wald F-statistic(Cragg & Donald, 1993), which is designed to detect whether the instrumental variables are weak instruments. If the correlation between the instrumental variable and the endogenous explanatory variable is weak, even if the instrumental variable satisfies the exogenous condition (unrelated to the error term), the estimated coefficients may still be inaccurate. The Wald F statistic is used to test whether this correlation is strong enough.

Kleibergen Paap Wald rk F-statistics is an improvement and extension of the Wald F statistic(Kleibergen & Paap, 2006), which considers the effects of heteroscedasticity and autocorrelation and is suitable for a wider range of scenarios.

4.4 The heterogeneous impact of different digital technologies

In order to investigate the differential impacts that different digital technologies have on the financial performance of enterprises, this study will redefine the categories of treatment and control groups. Those enterprises that have not adopted any digital technologies will be classified as the control group and assigned a numerical value of 0. For those enterprises that employ a specific digital technology, the numerical value will be assigned according to the corresponding technology name, with a value of 1.

As shown in Table 8 and Table 9, the impact of various digital technologies on the ROA and ROE of enterprises. The results indicate that adopting any digital technology can significantly enhance the ROA of enterprises, which is consistent with the findings in the existing literature. However, it should be noted that the regression coefficient of the impact of blockchain technology on ROA does not reach a significant level. This may be because although blockchain technology can enhance the security of enterprise information, its positive effects on stock prices often remain short-lived, while amplifying the risk of return of the enterprise. Furthermore, the current application of blockchain technology is still in its initial stage, and the market dominated by private chains and consortium chains, truly implementing a decentralized public chain is relatively rare.

Table 8 The impact of different digital technologies on ROA

	ROA_{AI}	ROA_{BD}	ROA_{CC}	ROA_{IoT}	ROA_{BC}	ROA_{MI}
Artificial Intelligence	0.612***					
	(0.172)					
Big Data		0.561***				
		(0.150)				
Cloud Computing			0.701***			
			(0.221)			
Internet to Thing				0.356**		
				(0.176)		
Blockchain					-0.122	
					(0.375)	
Mobile Internet						0.278**
						(0.142)
Model Selection	FE	FE	FE	FE	FE	FE
Observation	21487	18571	14549	22710	10924	17431
R^2	0.204	0.202	0.199	0.212	0.173	0.183

Table 9 The impact of different digital technologies on ROE

	ROE_{AI}	ROE_{BD}	ROE_{CC}	ROE_{IoT}	ROE_{BC}	ROE_{MI}
Artificial Intelligence	0.954***					
	(0.233)					
Big Data		1.336***				
		(0.345)				
Cloud Computing			1.117***			
			(0.437)			
Internet to Thing				0.778***		

				(0.231)		
Blockchain					-0.272	
					(0.456)	
Mobile Internet						0.644**
						(0.360)
Model Selection	FE	FE	FE	FE	FE	FE
Observation	21487	18571	14549	22710	10924	17431
R^2	0.133	0.129	0.155	0.118	0.088	0.142

4.5 Quantile regression

The results of the regression analysis reveal that the DT of businesses significantly enhances their financial performance. However, for companies with significant differences in financial performance, the impact of adopting digital technologies may vary. Additionally, the residuals of the models constructed using the least squares method are particularly sensitive to extreme values, which could lead to biased regression analysis results. In order to delve into the extent of the impact of DT on different financial performance companies, this section will employ the method of quantile regression. By selecting the five key quantiles of 10%, 25%, 50%, 75%, and 90%, a more detailed analysis of the relationship between DT and financial performance can be conducted.

The return on assets (ROA) is used as the dependent variable in the regression analysis, and the results of which are presented in Table 9. The results indicate that DT has a significant positive impact on the ROA of companies. However, as the quartile increases, the positive impact of DT on the ROA of companies at the 75% and 90% quartiles is not significant. This phenomenon may be explained by the fact that digital technologies, such as cloud computing, empower small and medium-sized enterprises in their growth phase. By implementing a light asset approach, these technologies effectively mitigate the inherent deficiencies in capital expenditure and scale expansion for such companies. Therefore, DT provides the possibility for underperforming companies to rapidly develop through technological innovation.

Table 9 The impact of DT on the regression results on ROA for different quantile levels.

Variables	QR_{90}	QR_{75}	QR_{50}	QR_{25}	QR_{10}
DT	0.1882	0.1571	0.3776*	0.1326*	0.3509**
	(0.262)	(0.134)	(0.223)	(0.073)	(0.155)
Controls	YES	YES	YES	YES	YES
Model Selection	FE	FE	FE	FE	FE
Observations	28041	28041	28041	28041	28041

In conjunction with the discussions in the preceding chapters, this study has compiled three primary conclusions. Firstly, the specific digital technologies adopted by businesses during the process of DT vary significantly, leading to substantial differences in their transformation outcomes. The current research data indicates that businesses that adopt blockchain technology do not exhibit significant improvements in financial performance. Secondly, the effectiveness of DT in businesses exhibits substantial heterogeneity. For businesses with poor financial performance, DT can bring about substantial positive effects. However, among businesses with superior financial performance, particularly those that are exceptional, the impact of DT is not evident. Finally, using different methodological approaches to construct DT indicators will lead to differences in the results of regression analysis. This point mirrors the debate in the introduction regarding the success or failure of DT in different businesses. This is because the paths, foundations, and measurement methods used by different businesses are unique, which results in diverse outcomes. This also implies that studying the DT of businesses requires differentiating between different digital technologies and financial foundations, and the necessity of constructing a unified and comparable corporate DT indicator.

4.6 Channel Analysis

Based on existing literature research, the path to DT can be summarized into three primary forms. Firstly, technology-driven or process optimization, which relies on introducing cutting-edge technologies to enhance the productivity of the company and subsequently improve its performance. Secondly, lean management, which focuses on process optimization and organizational changes to reduce production and operational costs, minimize waste from parts and raw materials, and enhance the financial performance of the company. Finally, customer experience optimization and marketing channel transformation, which emphasizes the optimization of the entire service process through accumulating customers and analyzing user consumption data to improve product loyalty, encourage repeat purchases, and increase revenues to enhance the financial performance of the company.

To validate the impact of production efficiency, the study selects total factor productivity (TFP) as the dependent variable. In calculating corporate TFP, solving the issue of endogeneity in the labor investment elasticity estimation using the OP method (1992) and LP method (2003) is crucial. The ACF method (2015) can effectively deal with the potential multiple collinearity problems in the estimation of labor investment elasticity using OP and LP methods.

Table 10 Channel Analysis

Variables	TFP_{sales}	TFP_{EVA}	Income	Cost	Cost/Income
DT	0.0188^{*}	0.0144^{**}	-6.42×10^{-4}	-0.0331^{***}	-0.0112^{**}
	(0.00761)	(0.00436)	(4.82×10^{-3})	(3.11×10^{-3})	(5.12×10^{-3})
Controls	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES	YES
Count	24503	20079	28041	28041	28041
R^2	0.0662	0.377	0.642	0.763	0.0883

Note: The table adopts the ACF method and calculates two forms of TFP based on sales revenue and economic value added. Calculation method using Cobb-Douglas production function: $TFP=\frac{Y}{K^{\alpha}L^{1-\alpha}}$, with Y represent the output value adjusted for sales revenue, K for capital input, L for labor input，and α is the output elasticity of capital. Y could be adjusted to economic value added.

To investigate the pathways of income growth, this study employs the natural logarithm of total income as the dependent variable in Table 10. Simultaneously, in order to examine the path of cost reduction, this table also sets the natural logarithm of total cost as the explanatory variable. Further, by introducing the Cost/Income ratio as another explanatory variable, this study aims to comprehensively evaluate the financial performance of enterprises from the dimensions of both cost and income. The findings reveal that, after the execution of DT, the TFP of enterprises significantly increases, while total cost significantly decreases. These results validate the existence of cost saving pathways and efficiency improvement paths. However, there is no significant increase in total income, suggesting that the pathways of income growth have not been empirically supported. Moreover, the Cost/Income result indicates that the cost per unit of revenue decreases, indicating that the cost-income ratio decreases. This implies that, after enterprises adopt digital technologies, their financial performance overall improves.

5. Conclusion and Implication

5.1 Conclusion

In the context of the rapid development of the global digital economy, the substantial progress of enterprise DT has become an important component of the micro-foundation. Simultaneously, academic attention on the practice of enterprise DT has also intensified. However, the main problems in measuring enterprise DT indicators currently include the lack of unified and explicit object and the lack of scientific and precise method. This has led to serious discrepancies in understanding the status of enterprise DT and its effects. Based on the cutting-edge language model development, we have established a new set of enterprise DT indicators in this paper. Multiple cross-validation results indicate that our newly developed indicators are significantly superior to the dominant word dictionary method and are consistent with the mainstream practices of enterprise DT.

In the later parts of this paper, we demonstrate through novel methods that American publicly listed companies have significantly improved their financial performance through enterprise DT. For those companies with relatively poor financial performance, the positive impact of DT is particularly prominent. Further analysis reveals that the primary mechanism by which DT promotes financial performance growth lies in enhancing operating efficiency and reducing operational costs.

5.2 Implication for Future Study

The findings of this study hold significant policy implications for promoting the digitization process in enterprises and achieving high-quality economic growth. Firstly, the national government, the business community, and academia should collaborate to provide necessary support and fundamental research for the theoretical framework and practical applications of enterprise digitization. Facing the major challenges encountered by enterprises in the process of DT, one of the fundamental reasons is the lack of a unified and authoritative measurement standard for DT and a profound analysis of its effects. This situation seriously hinders the timely tracking and summation analysis by government agencies of the success or failure of corporate DT. To construct a comprehensive, unified, and precise DT indicator system and database, detailed information is required from various dimensions such as digital technology application, human resource management, and financial investment. However, current academic research in this field is largely reliant on text analysis methods. This study, employing large language models, constructed a DT indicator system, providing a preliminary foundation for research on DT and pointing to feasible research directions. However, the validity of this outcome still needs to be confirmed through the mutual verification of relevant data.

Secondly, companies should persistently pursue DT. DT, by enhancing efficiency and reducing costs, contributes to improving the financial performance of enterprises. Particularly in the context of globalization, enterprises in developing countries face traditional advantages of multinational corporations in terms of scale, brand, and capital. Therefore, they should leverage the post-industrial advantages provided by the digital economy era, and strengthen their efforts to promote DT in their

businesses to enhance their competitiveness in the international market.
Thirdly, governments should implement differentiated DT incentives according to the specific situations of enterprises. The results of this study indicate that DT has a significant impact on the financial performance of both average and poor-performing enterprises, but its effect on exceptional-performing enterprises is not significant.

Reference

Ackerberg, D. A., Caves, K., & Frazer, G. (2015). Identification Properties of Recent Production Function Estimators. *Econometrica*, *83*(6), 2411–2451. https://doi.org/10.3982/ECTA13408

Agostino, D., & Costantini, C. (2021). A measurement framework for assessing the digital transformation of cultural institutions: The Italian case. *Meditari Accountancy Research*, *30*(4), 1141–1168. https://doi.org/10.1108/MEDAR-02-2021-1207

Appio, F. P., Frattini, F., Petruzzelli, A. M., & Neirotti, P. (2021). Digital Transformation and Innovation Management: A Synthesis of Existing Research and an Agenda for Future Studies. *Journal of Product Innovation Management*, *38*(1), 4–20. https://doi.org/10.1111/jpim.12562

Araci, D. (2019). *FinBERT: Financial Sentiment Analysis with Pre-trained Language Models* (arXiv:1908.10063). arXiv. https://doi.org/10.48550/arXiv.1908.10063

Asikpo, N. A. (2024). Impact of Digital Transformation on Financial Reporting in the 21st Century. *International Journal of Comparative Studies and Smart Education*, *1*(1), Article 1.

Babina, T., Fedyk, A., He, A., & Hodson, J. (2024). Artificial intelligence, firm growth, and product innovation. *Journal of Financial Economics*, *151*, 103745. https://doi.org/10.1016/j.jfineco.2023.103745

Berger, E. D. (2020). *CSRankings*. https://csrankings.org

Brynjolfsson, E., Li, D., & Raymond, L. R. (2023). *Generative AI at Work* (Working Paper 31161). National Bureau of Economic Research. https://doi.org/10.3386/w31161

Chawla, R. N., & Goyal, P. (2021). Emerging trends in digital transformation: A bibliometric analysis. *Benchmarking: An International Journal*, *29*(4), 1069–1112. https://doi.org/10.1108/BIJ-01-2021-0009

China Academy of Information and Communications Technology. (2024). *White Paper on Global Digital Economy (2023 Edition)*. http://www.caict.ac.cn/kxyj/qwfb/bps/202401/P020240326601000238100.pdf

Cragg, J. G., & Donald, S. G. (1993). Testing Identifiability and Specification in Instrumental Variable Models. *Econometric Theory*, *9*(2), 222–240. https://doi.org/10.1017/S0266466600007519

Davenport, T. H., & Westerman, G. (2018). *Why So Many High-Profile Digital Transformations Fail*.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding* (arXiv:1810.04805). arXiv. https://doi.org/10.48550/arXiv.1810.04805

Drechsler, K., Gregory, R., Wagner, H.-T., & Tumbas, S. (2020). At the Crossroads between Digital Innovation and Digital Transformation. *Communications of the Association for Information Systems*, *47*(1). https://doi.org/10.17705/1CAIS.04723

George, G., & Schillebeeckx, S. J. D. (2022). Digital transformation, sustainability, and purpose in the multinational enterprise. *Journal of World Business*, *57*(3), 101326. https://doi.org/10.1016/j.jwb.2022.101326

Ghobakhloo, M. (2020). Industry 4.0, digitization, and opportunities for sustainability. *Journal of Cleaner Production*, *252*, 119869. https://doi.org/10.1016/j.jclepro.2019.119869

Guandalini, I. (2022). Sustainability through digital transformation: A systematic literature review for research guidance. *Journal of Business Research*, *148*, 456–471. https://doi.org/10.1016/j.jbusres.2022.05.003

Guo, X., Li, M., Wang, Y., & Mardani, A. (2023). Does digital transformation improve the firm's performance? From the perspective of digitalization paradox and managerial myopia. *Journal of Business Research*, *163*, 113868. https://doi.org/10.1016/j.jbusres.2023.113868

Hu, E. J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L., & Chen, W. (2021). *LoRA: Low-Rank Adaptation of Large Language Models* (arXiv:2106.09685). arXiv. http://arxiv.org/abs/2106.09685

Huang, A. H., Wang, H., & Yang, Y. (2023). FinBERT: A Large Language Model for Extracting Information from Financial Text*. *Contemporary Accounting Research*, *40*(2), 806–841. https://doi.org/10.1111/1911-3846.12832

Huang, W., Ma, X., Qin, H., Zheng, X., Lv, C., Chen, H., Luo, J., Qi, X., Liu, X., & Magno, M. (2024). *How Good Are Low-bit Quantized LLaMA3 Models? An Empirical Study* (arXiv:2404.14047). arXiv. https://doi.org/10.48550/arXiv.2404.14047

İnel, M. (2019). An empirical study on measurement of efficiency of digital transformation by using data envelopment analysis. *Management Science Letters*, *9*(4), 549–556.

Jacques Bughin, Jonathan Deakin, & Barbara O'Beirne. (2019). *Digital transformation: Improving the odds of success*. McKinsey. https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/McKinsey%20Di gital/Our%20Insights/Digital%20transformation%20Improving%20the%20odds%20of%20success/Digital-transformation-Improving-the-odds-of-success-final.pdf

Kleibergen, F., & Paap, R. (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics*, *133*(1), 97–126. https://doi.org/10.1016/j.jeconom.2005.02.011

Kraus, S., Durst, S., Ferreira, J. J., Veiga, P., Kailer, N., & Weinmann, A. (2022). Digital transformation in business and management research: An overview of the current status quo. *International Journal of Information Management*, *63*, 102466. https://doi.org/10.1016/j.ijinfomgt.2021.102466

Larysa Yakymova, Andrea Novotná, & Vasyl Kuz. (2022). MEASURING INDUSTRY DIGITAL TRANSFORMATION WITH A COMPOSITE INDICATOR: A CASE STUDY OF THE UTILITY INDUSTRY. *Journal of International Studies*, *15*(1), 168–180.

Levinsohn, J., & Petrin, A. (2003). Estimating Production Functions Using Inputs to Control for Unobservables. *The Review of Economic Studies*, *70*(2), 317–341. https://doi.org/10.1111/1467-937X.00246

Li, J., Xia, J., & Zajac, E. J. (2018). On the duality of political and economic stakeholder influence on firm innovation performance: Theory and evidence from Chinese firms. *Strategic Management Journal*, *39*(1), 193–216. https://doi.org/10.1002/smj.2697

Li, S., Gao, L., Han, C., Gupta, B., Alhalabi, W., & Almakdi, S. (2023). Exploring the effect of digital transformation on Firms' innovation performance. *Journal of Innovation & Knowledge*, *8*(1), 100317. https://doi.org/10.1016/j.jik.2023.100317

Meta. (n.d.). *Introducing Meta Llama 3: The most capable openly available LLM to date*. Meta AI.

Retrieved 9 June 2024, from https://ai.meta.com/blog/meta-llama-3/

Mielli, F., & Bulanda, N. (2019). Digital Transformation: Why Projects Fail, Potential Best Practices and Successful Initiatives. *2019 IEEE-IAS/PCA Cement Industry Conference (IAS/PCA)*, 1–6. https://doi.org/10.1109/CITCON.2019.8729105

Nambisan, S., Wright, M., & Feldman, M. (2019). The digital transformation of innovation and entrepreneurship: Progress, challenges and key themes. *Research Policy*, *48*(8), 103773. https://doi.org/10.1016/j.respol.2019.03.018

North American Industry Classification System (NAICS) U.S. Census Bureau. (n.d.). Retrieved 13 June 2024, from https://www.census.gov/naics/

Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, *381*(6654), 187–192. https://doi.org/10.1126/science.adh2586

Nwankpa, J. K., & Datta, P. (2017). Balancing exploration and exploitation of IT resources: The influence of Digital Business Intensity on perceived organizational performance. *European Journal of Information Systems*, *26*(5), 469–488. https://doi.org/10.1057/s41303-017-0049-y

Olley, G. S., & Pakes, A. (1992). *The Dynamics of Productivity in the Telecommunications Equipment Industry* (Working Paper 3977). National Bureau of Economic Research. https://doi.org/10.3386/w3977

Oludapo, S., Carroll, N., & Helfert, M. (2024). Why do so many digital transformations fail? A bibliometric analysis and future research agenda. *Journal of Business Research*, *174*, 114528. https://doi.org/10.1016/j.jbusres.2024.114528

Pappas, I. O., Mikalef, P., Giannakos, M. N., Krogstie, J., & Lekakos, G. (2018). Big data and business analytics ecosystems: Paving the way towards digital transformation and sustainable societies. *Information Systems and E-Business Management*, *16*(3), 479–491. https://doi.org/10.1007/s10257-018-0377-z

Pereira, V., Hadjielias, E., Christofi, M., & Vrontis, D. (2023). A systematic literature review on the impact of artificial intelligence on workplace outcomes: A multi-process perspective. *Human Resource Management Review*, *33*(1), 100857. https://doi.org/10.1016/j.hrmr.2021.100857

Pramanik, H. S., Kirtania, M., & Pani, A. K. (2019). Essence of digital transformation—Manifestations at large financial institutions from North America. *Future Generation Computer Systems*, *95*, 323–343. https://doi.org/10.1016/j.future.2018.12.003

Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (n.d.). *Improving Language Understanding by Generative Pre-Training*.

Sun, Y., Wang, S., Li, Y., Feng, S., Tian, H., Wu, H., & Wang, H. (2020). ERNIE 2.0: A Continual Pre-Training Framework for Language Understanding. *Proceedings of the AAAI Conference on Artificial Intelligence*, *34*(05), Article 05. https://doi.org/10.1609/aaai.v34i05.6428

Tavana, M., Shaabani, A., Raeesi Vanani, I., & Kumar Gangadhari, R. (2022). A Review of Digital Transformation on Supply Chain Process Management Using Text Mining. *Processes*, *10*(5), Article 5. https://doi.org/10.3390/pr10050842

The World Bank. (2023). *Digital Progress and Trends Report*. Digital Progress and Trends Report 2023. https://www.worldbank.org/en/publication/digital-progress-and-trends-report?continueFlag=2bbd9c1c0494336106d9ddae888d1220

Tong, H., Li, J., Wu, N., Gong, M., Zhang, D., & Zhang, Q. (2024). *Ploutos: Towards interpretable stock movement prediction with financial large language model* (arXiv:2403.00782). arXiv. https://doi.org/10.48550/arXiv.2403.00782

Vial, G. (2019). Understanding digital transformation: A review and a research agenda. *The Journal of Strategic Information Systems*, *28*(2), 118–144. https://doi.org/10.1016/j.jsis.2019.01.003

Wang, X., Li, Y., Tian, L., & Hou, Y. (2023). Government digital initiatives and firm digital innovation: Evidence from China. *Technovation*, *119*, 102545. https://doi.org/10.1016/j.technovation.2022.102545

WIPO. (2022). *World Intellectual Property Report 2022:* Unknown. https://doi.org/10.34667/TIND.45356

World Bank. (2024). *Digital Progress and Trends Report 2023*. The World Bank. https://doi.org/10.1596/978-1-4648-2049-6

Xie, X., & Wang, S. (2023). Digital transformation of commercial banks in China: Measurement, progress and impact. *China Economic Quarterly International*, *3*(1), 35–45. https://doi.org/10.1016/j.ceqi.2023.03.002

Yang, Y., UY, M. C. S., & Huang, A. (2020). *FinBERT: A Pretrained Language Model for Financial Communications* (arXiv:2006.08097). arXiv. https://doi.org/10.48550/arXiv.2006.08097

Yifeng, P. (2024). *Internet sentiment exacerbates intraday overtrading, evidence from A-Share market* (arXiv:2404.12001). arXiv. https://doi.org/10.48550/arXiv.2404.12001

Zhai, H., Yang, M., & Chan, K. C. (2022). Does digital transformation enhance a firm's performance?

Evidence from China. *Technology in Society*, *68*, 101841. https://doi.org/10.1016/j.techsoc.2021.101841

Zhang, B., Yang, H., Zhou, T., Ali Babar, M., & Liu, X.-Y. (2023). Enhancing Financial Sentiment Analysis via Retrieval Augmented Large Language Models. *Proceedings of the Fourth ACM International Conference on AI in Finance*, 349–356. https://doi.org/10.1145/3604237.3626866

Zhang, Y., & Zhang, H. (2023). FinBERT–MRC: Financial Named Entity Recognition Using BERT Under the Machine Reading Comprehension Paradigm. *Neural Processing Letters*, *55*(6), 7393–7413. https://doi.org/10.1007/s11063-023-11266-5

Zolas, N., Kroff, Z., Brynjolfsson, E., McElheran, K., Beede, D. N., Buffington, C., Goldschlag, N., Foster, L., & Dinlersoz, E. (2020). *Advanced Technologies Adoption and Use by U.S. Firms: Evidence from the Annual Business Survey* (Working Paper 28290). National Bureau of Economic Research. https://doi.org/10.3386/w28290