Exploring the Digital Leverage: How Big Data Technologies Computationally Drive Firm Performance Progress

Ying Dong[1*], *Sun Wei*[2]

[1] *College of Tourism and Culture, Liaodong University, China*
[2] *Liaoning Siji Technology Service Co., Ltd., China*
dy8288@126.com, lnddsw@126.com

Abstract

In the era of rapid technological advancement, advancements in computer hardware, such as high-performance processors and vast storage capabilities, have laid the foundation for handling complex data analytics tasks. Concurrently, breakthroughs in software algorithms and machine learning models have revolutionized the way businesses process and interpret data. These computer technologies, particularly big data analytics, have emerged as transformative forces reshaping business landscapes. They enable firms to process vast amounts of data, uncover actionable insights, and optimize decision-making processes. Methodologically, this study employs Python for keyword frequency statistics to quantify the presence and impact of big data technologies within firm-related data sources. By analyzing the frequency of relevant keywords, we gain a nuanced understanding of the extent to which big data technologies are integrated and utilized by firms. Furthermore, multiple regression analysis is utilized to conduct a mechanism analysis, exploring the relationships between big data technologies, firm performance, and other influencing factors. Our findings reveal that big data technologies significantly drive firm performance development by enhancing operational efficiency, enabling data-driven decision-making, and fostering innovation. Notably, a reduced supply chain concentration acts as a mediator, amplifying the positive impact of big data technologies on firm performance. Additionally, top management team rejuvenation positively moderates the effects of big data technologies on firm performance. This suggests that as firms leverage big data technologies, they are able to streamline their supply chains, leading to improved performance outcomes. Additionally, the age of the top management team moderates this relationship, with younger teams being more receptive and effective in leveraging big data technologies to drive performance.

Keywords: Big data technologies, Firm performance development, Supply chain concentration, Top management team age

1 Introduction

Big data generally refers to data sets that are beyond the capabilities of traditional database tools in terms of acquisition, storage, management, and analysis [1-2]; it has the characteristics of massive data scale, rapid data flow, diverse data types and low value density. From the initial application of Hadoop technology in the Internet, telecommunications, finance and other industries, to the active layout of big data tool chain by various cloud manufacturers, around Hadoop, MPP database, agile BI and other key technologies, a number of start-ups were born and promoted the penetration of big data applications from consumer Internet to manufacturing, agriculture, energy, retail and other industrial Internet. The industrial chain of big data industry can be divided into three main links: upstream, midstream and downstream. The upstream of the big data industry is mainly concentrated in the basic support layer, including network equipment, computer equipment, storage equipment and other hardware supply. The downstream of the big data industry is the big data application market, which is an important link in the big data industry chain to realize value. With the continuous improvement of big data research technology level, big data has been widely used in government, industry, finance, transportation, telecommunications and spatial geography and other industries. In the current global and technology-driven business environment, the application of big data technology (BDT) has become a key tool for enterprises to gain a competitive advantage [3-5]. By providing in-depth market insights and enhancing decision-making support, BDT enables enterprises to remain competitive under complex and changing market conditions.

Through literature review, it is found that enterprise performance has always been one of the hot spots of scholars' research. The research results show that there are many internal and external factors affecting enterprise performance. Most scholars conduct research from the aspects of enterprise digital transformation and artificial intelligence application, and big data technology itself is one of the internal factors of enterprises. At present, there are few literature on how big data technology affects enterprise performance, so how big data technology affects enterprise performance is the focus of this paper. In addition, at present, there are few literature that systematically analyze the internal mechanism of big data technology application affecting enterprise performance. Supply chain concentration is an important indicator

*Corresponding Author: Ying Dong; Email: dy8288@126.com
DOI: https://doi.org/10.70003/160792642026052703013

to measure the dependence of enterprises in the supply chain. Recent studies have shown that by applying BDT, enterprises can optimize their supply chain management and reduce their dependence on a few suppliers, thus improving the resilience and efficiency of the supply chain [6-7]. In addition, BDT can help enterprises find new business opportunities around the world and improve performance through better market positioning and customer service [8]. At the same time, the age structure of the top management team also has a decisive impact on the effectiveness of enterprises using BDT. The average age of the top management team has a negative impact on the digital orientation of firms [9]. That means younger management teams have a preference for high-risk decisions and expect to be able to achieve high returns through means such as rapid expansion, so they are more likely to embrace new technologies and innovations, which are critical to keeping a business competitive in a changing technological environment. In view of this, on the basis of existing studies, this paper explores the mechanism of big data technology application affecting enterprise performance from the perspective of supply chain concentration and uses the data of listed companies from 2013 to 2022 for empirical testing. At the same time, this study explores how the age of the senior management team adjusts the impact of BDT on the performance development of enterprises in order to provide guidance for the human resources and technology strategy of enterprises. This study will focus on how BDT affects the development of enterprise performance through the intermediary role of supply chain concentration and the adjustment of the age of senior management teams.

To summarize, this study aims to deeply explore the impact of BDT on enterprise performance in different enterprise environments through theoretical and empirical analysis. By revealing these complex relationships, this study hopes to provide theoretical basis and practical suggestions for enterprises to implement effective big data strategies.

2 Literature Review and Hypothesis Development

2.1 Big Data Technologies and Firm Performance Development

With the rapid development of big data technology (BDT), the multi-faceted impact on enterprise performance has attracted more and more attention. This section explores the impact of BDT on firm performance development based on empirical research and theoretical frameworks.

Enterprises apply big data technology mainly to acquire, store, process and analyze big data, and conduct data mining, data analysis and data visualization operations for massive data, focusing on discovering hidden opportunities and mining potential value in massive data, with the ultimate goal of improving the efficiency of production and operation decision-making and creating greater value and benefits. The application of big data technology by enterprises can promote sales, revenue, the number of employees, net profit margin, innovative management, the development of new products and new services, as well as the adoption of new information technologies [10], all of these can promote the flow of data resources, improve the utilization efficiency, and promote the improvement of enterprise performance by giving full play to the potential of data elements. First, big data technology helps enterprises to screen customers, carry out precision marketing, and increase sales. Precision marketing is one of the significant advantages of big data applications. By deeply analyzing consumer behavior data, companies can design more personalized products and services that increase customer satisfaction and loyalty. Research by Davenport and Harris [11] and McAfee and Brynjolfsson [12] points out that the use of big data for market analysis and consumer insight can significantly improve marketing effectiveness and thus promote sales growth. Second, big data technology can help enterprises reduce production costs. The reduction of production cost is directly related to the profit rate of enterprises. Ge and Jackson [13] found through their research on the automotive industry that big data technology can save a lot of costs for enterprises by optimizing production processes and improving energy efficiency. Real-time data analysis enables companies to accurately forecast demand, optimize inventory management, reduce inventory overhang and overproduction, and effectively control costs. Third, big data technology can help identify business opportunities. The application of big data technology is not limited to the optimization of existing businesses, but more importantly, it can reveal new business opportunities. Bughin et al. [8] emphasize that by analyzing large amounts of market and consumer data, companies can identify unmet needs, explore new market potential, and drive innovation in products and services. Fourth, big data technology can help enterprises optimize management. Management optimization is another important benefit of big data technology. Wamba et al. [14] demonstrated the role of big data technology in improving decision quality, optimizing resource allocation, and enhancing operational efficiency. Especially in the fields of supply chain management, customer relationship management and human resource management, big data technology provides more accurate and real-time information support to help managers make more effective decisions. Based on the above discussion, we can draw the following hypothesis:

H1: The adoption and integration of big data technology significantly promotes the development of enterprise performance.

2.2 The Mediating Effect of Supply Chain Concentration

2.2.1 Big Data Technologies and Supply Chain Concentration

With the in-depth development of digital transformation, the application of big data technology in supply chain management has become increasingly important. This technology not only changes the operation mode of the supply chain but also has a significant impact on the concentration of the supply chain. Big data technology has a driving effect on enterprise supply management. With the deepening of digital transformation, the application of big data technology in the field of supply chain management has become a trend. By integrating and analyzing massive amounts of data, these technologies help companies improve the accuracy of market demand forecasts, optimize inventory management, and increase the flexibility and efficiency of supply chains. The research of Xu et al. [15] shows that digital transformation, especially the application of big data technology, can significantly reduce the concentration of supply chain. Big data technology can help businesses identify new suppliers and customers globally, improve supply chain flexibility and adaptability, and reduce the risk of market volatility or supply chain disruptions. Jens et al. [16] explores the application of big data in optimizing supply chain management. These technologies not only improve the efficiency and reliability of the supply chain but also enable the supply chain to respond more flexibly to market changes through real-time monitoring and data analysis, further contributing to the reduction of supply chain concentration. Through the above analysis, it can be seen that big data technology plays a key role in reducing the concentration of supply chains. The application of this technology not only promotes the diversification and global management of the supply chain but also improves the transparency and efficiency of the supply chain. Therefore, we propose the following hypothesis:

H2: Big data technology reduces supply chain concentration.

2.2.2 Supply Chain Concentration and Firm Performance Development

In the era of digital economy, global supply chain management has become the key for enterprises to gain competitive advantage. Through global procurement, enterprises can reduce costs, while global sales help to expand the market, both of which have a direct impact on enterprise performance. However, the increase in supply chain concentration, while bringing cost advantages, may also increase the risk of supply chain disruption. The introduction of big data technology is regarded as a key means to solve this contradiction. First, big data technology can improve enterprise supply chain resilience. In the tide of globalization, enterprises choose low-cost suppliers to save costs through global procurement strategy. Research by Xu et al. [15] reveals the key role of big data technology in optimizing supply chain management and reducing concentration, showing that big data can significantly improve the market adaptability and cost efficiency of the procurement process. In addition, the analysis and application of big data helps enterprises evaluate and select suppliers more accurately, avoiding over-reliance on a few suppliers and improving the resilience of the entire supply chain. Second, big data technology can expand the global sales network of enterprises. By leveraging big data technology, enterprises can enhance the efficiency of their supply chains and their ability to accurately match suppliers. At the same time, they can accurately analyze market trends, identify new sales opportunities, and respond more effectively to consumer demands. All these are conducive to the company expanding its sales network and increasing its market share globally, which is helpful to improve the enterprise's sales performance and market competitiveness. Third, big data technology can improve enterprise communication efficiency and cost optimization. The optimization of communication cost is another key point of supply chain management. Liu [17] mentioned in the research that the application of big data technology is of great significance to the management and planning of the supply chain. It can achieve product traceability and precise marketing, improve the response speed and collaboration ability of the supply chain, realize intelligent storage and distribution, and enhance the terminal delivery efficiency, thereby maximizing the profit of the supply chain. Based on the above literature, we propose:

H3: excessive concentration of supply chain may adversely affect enterprise performance, especially in the face of the risk of supply chain disruption.

H4: big data technology can effectively promote the improvement of enterprise performance by dispersing the concentration of supply chain.

2.3 The Moderating Effect of Top Management Team Age

The high-level echelon theory clearly points out that the individual characteristics of the members of the senior management team, such as age, educational background, professional experience, etc., have a significant impact on the outcome of team decision-making. These characteristics not only shape the way of thinking of team members, but also affect their decision-making preferences, which in turn lead to the adjustment of the overall performance and strategic direction of the organization. Therefore, a deep understanding of the characteristics of the members of the senior management team is of great significance for predicting and guiding the development of the organization. The age structure of the top management team is an important factor of the enterprise's strategy execution and innovation ability. Younger management teams are often seen as having a broader vision, broader knowledge, and greater mastery of digital technology. These qualities are particularly important for developing incremental innovation in the high-tech industry. First, younger management teams are more sensitive to new technologies in the industry. The technological sensitivity of young managers may have promoted innovative activities in the field of big data and related technologies. They tend to pay more attention to the latest industry trends and seek to achieve business model breakthroughs through technology. In Blank's [18] study, younger R&D teams in the high-tech industry showed greater ability to innovate, suggesting that younger management teams may be more inclined to adopt advanced technologies such as big data to drive business growth. Second, younger management teams are faster to master industry technology and integrate into strategic decisions. The information technology proficiency of young managers has a direct impact on strategic decision-making. Younger executives are more likely to embrace and use big data analytics to optimize the decision-making process, enabling companies to make more accurate business decisions based on deep data insights. Third, younger management teams have more advantages in knowledge updating and cross-border integration. Another advantage of young managers is their initiative in updating knowledge. Tushman and O'Reilly [19] emphasized the role of cross-border knowledge integration in promoting innovation, and young managers are often more flexible in integrating new knowledge from different disciplines and fields to promote cross-border innovation. Fourth, younger management teams are more open and digitally adaptable. Young senior management teams often excel in openness to the outside world and digital adaptability. Young managers are often open to new ideas and new ways of working, making them more comfortable introducing and promoting new technologies. This mindset not only helps companies adapt quickly to digital transformation but also provides a strong support for companies to explore new business opportunities. To summarize, we assume that:

H5: The impact of senior managers' age on big data technology on the development of enterprise performance. That is, the aging of senior managers will weaken the positive role of big data technology in the development of enterprise performance.

3 Methodology

3.1 Selection of Samples and Data Sources

This study focuses on Chinese listed companies as the subjects of analysis. Given the availability of data, this study establishes the research period from 2013 to 2022. Data sources comprise of the CSMAR database in China. During data processing, ST stocks, financial industry and samples with data missing are excluded. Secondly, Winsorize 1% and 99% quartile of continuous variables to ensure that the data is real and reasonable.

Table 1. Steps of keywords frequency statistics

Step1. Text Data Acquisition

First, it is necessary to obtain the text data of the listed company's reports. This data usually exists in formats such as PDF, HTML, or plain text. Libraries like *requests* can be used for web scraping, and *PyPDF2* or *pdfplumber* can be used to parse PDF files.

Step2. Text Preprocessing

Before counting word frequencies, the text needs to be preprocessed to improve the accuracy of the analysis. The main steps include:

Noise Removal: Remove irrelevant information such as HTML tags, special characters, headers, and footers from the text.

Tokenization: Split the text into words or phrases. For Chinese text, the jieba library can be used for tokenization.

Stop Words Removal: Remove common meaningless words (such as "and", "is" in Chinese). This can be achieved by maintaining a stop words list.

Lemmatization/Stemming: For English text, lemmatization or stemming can be performed to unify different forms of the same word.

Step3. Word Frequency Counting

The core of word frequency counting is to calculate how many times each keyword appears in the text. The formula is:

Word Frequency(w)= Number of times keyword w appears/Total number of words in the text

Or simply use the count of occurrences as the measure of word frequency, depending on the analysis requirements.

In Python, the *collections. Counter* can be used to conveniently count word frequencies. For example:

```
Python
from collections import Counter
import jieba
# Sample text
text = " This is a sample text to demonstrate the method of word frequency statistics."
# Word segmentation,
words = jieba.lcut(text)
# Word frequency statistics
word_counts = Counter(words)
print(word_counts)
```

Step4. Keyword Extraction and Matching

To count the word frequencies of core keywords, a list of keywords needs to be defined in advance. Then, iterate through the tokenized text and count the occurrences of these keywords.

```
Python
keywords = ["Sample text", "Word segmentation", "Word frequency statistics"]
keyword_counts = {keyword: 0 for keyword in keywords}
for word in words:
    if word in keywords:
        keyword_counts[word] += 1
print(keyword_counts)
```

Step5. Data Analysis and Visualization

After counting the word frequencies, further data analysis can be performed, such as calculating TF-IDF (Term Frequency-Inverse Document Frequency) to evaluate the importance of keywords. Additionally, libraries like *matplotlib* or *seaborn* can be used for visualization to present the results more intuitively.

3.2 Definition of Variables

3.2.1 Explanatory Variable: Big Data Technology

Referring to Zhang et al. [20], this study uses word frequency statistics and text analysis to measure digital technology application. Specifically, this study employs machine learning techniques and Python software to extract annual reports from publicly traded tourism companies. Subsequently, it performs keyword frequency analysis on terms associated with digital technology application. Referring to Zhang et al. [20], the keywords related to big data technology applications include "big data", "virtual reality", "data visualization" "text mining", "data mining", "credit reporting" etc. Then, the natural logarithm of the total word frequency plus one is applied for data processing.

To guarantee the study's quality, the steps of keywords frequency statistics can be seen in Table 1.

3.2.2 Explained Variable: Enterprise Performance

The financial metrics of ratio of net profit to operating income is employed to assess the performance of enterprises. In assessing corporate performance, the net profit margin, which compares net profit to operating income, offers a direct measure of a company's profitability. Unlike the debt-to-asset ratio, which focuses on solvency by evaluating liabilities against assets, or return on equity (ROE), which gauges returns for shareholders, the net profit margin specifically highlights how efficiently a company converts sales into profit. This metric strips away complexities related to capital structure, providing a clear view of operational profitability. Its simplicity and focus on core business performance make it an essential tool for analysts and investors seeking to evaluate and compare companies effectively.

3.2.3 Mediating Variable: Supply Chain Concentration

In this study, the metric used for gauging supplier concentration (SCC) is determined by dividing the purchasing volumes from the top five suppliers by the organization's cumulative purchase volume.

3.2.4 Moderating Variable: Top Management Team Age (TMTage)

There are various methods to measure the age heterogeneity of the senior management team. Here, the average age of directors and supervisors is used to measure the variable.

3.2.5 Measurements of the Control Variables

To ensure a comprehensive assessment that negates extraneous in fluences, we incorporate controls for various firm-specific variables. To further enhance the model's precision, time and industry dummy variables are introduced to fix the model bidirectionally. The measurements of the variables are described in Table 2.

Table 2. Descriptions of variables

Classification	Variable	Abbreviation	Measurement
Dependent variable	Firm Performance	FP	(Income - Costs) / Income
	Firm Performance_sub	FP_sub	Net Profit / Operating Income
Independent variable	Big Data Technologies	BD	Ln (big data technology word frequency statistics +1)
Mediating variable	Supply Chain	SCC	(Proportion of purchases from top 5 suppliers + proportion of sales to top 5 customers) / 2
Moderating variable	TMT Age	TMTAge	Average age of TMT
Control variable	Firm Size	Size	Natural logarithm of annual total assets
	Firm Age	FirmAge	Ln (current year - year the company was founded +1)
	Asset-liability ratio	Lev	Total liabilities/ Total assets
	Net profit margin on total assets	ROA	Net profit / Average balance of total assets
	Investment Level	Invest	Cash paid for construction of fixed assets, intangible assets and other long-term assets / Total assets
	Current Ratio	Liquid	Current Assets / Current Liabilities
	Whether state-owned enterprises	SOE	1 for state-controlled enterprises, 0 for others

3.3 Model Construction

For the purpose of this study, a foundational regression model is constructed as demonstrated in the following equation:

$$FT_{i,t} = \beta_0 + \beta_1 BD_{i,t} + \beta_n C_{i,t} + Year_t + Industry_i + \varepsilon_{i,t} \quad (1)$$

$$SCC_{i,t} = \beta_0 + \beta_1 BD_{i,t} + \beta_n C_{i,t} + Year_t + Industry_i + \varepsilon_{i,t} \quad (2)$$

$$FT_{i,t} = \beta_0 + \beta_1 BD_{i,t} + \beta_2 SCC_{i,t} + \beta_n C_{i,t} + Year_t + Industry_i + \varepsilon_{i,t} \quad (3)$$

$$FT_{i,t} = \beta_0 + \beta_1 TMTage_{i,t} + \beta_2 BD_{i,t} + \beta_3 TMTage_{i,t} * BD_{i,t} + \beta_n C_{i,t} + Year_t + Industry_i + \varepsilon_{i,t} \quad (4)$$

In this equation, 'i' denotes the company, while 't' signifies the year. The symbol 'β' stands for the corresponding regression coefficients. The term 'C' encompasses all control variables. The placeholders 'Year' and 'Industry' denote the fixed influences of respective years and industries. Here, 'BD' designates the intensity of Big Data Technology, 'FP' indicates the measure of firm performance, 'SCC' indicates the measure of supply chain concentration, 'TMTAge' represents average age of directors and supervisors, and 'ε' captures the error term of the model.

4 Empirical Results

4.1 Descriptive Statistics

Table 3 illustrates descriptive statistics results of the research data.

Table 3. Summary statistics results

Variables	Observation	Mean	SD	Min	Median	Max
FP	38228	0.282	0.175	-0.021	0.249	0.830
FP_sub	38228	0.059	0.190	-1.036	0.066	0.489
BD	38228	0.455	0.842	0.000	0.000	3.738
SCC	38228	0.302	0.174	0.023	0.278	0.806
TMTAge	38228	49.105	3.206	41.330	49.150	56.620
Size	38228	22.161	1.288	19.793	21.975	26.139
FirmAge	38228	2.880	0.350	1.792	2.944	3.526
Lev	38228	0.432	0.206	0.056	0.426	0.906
ROA	38228	0.039	0.067	-0.238	0.038	0.228
Invest	38228	0.062	0.067	0.000	0.040	0.362
Liquid	38228	2.387	2.378	0.314	1.630	15.413
SOE	38228	0.369	0.482	0.000	0.000	1.000

4.2 Benchmark Regression

Table 4 sheds light on the findings from the baseline regression analysis. Column (1) shows the baseline model without controls. Columns (2) adds industry and time fixed effects, as well as control variables. Delving deeper into the exhaustive model outlined in Column (2), a unit escalation in big data technology corresponds to a 0.010-unit surge in firm Performance metrics, a finding significant at the 1 % threshold. H1 is supported here.

4.3 Mechanism Regression

According to the data in the first column of Table 5, without the introduction of any control variables, big data technology significantly negatively affects the concentration degree of the supply chain, and the regression coefficient is -0.020, which is significant at the level of 1%. However, after adding many control variables, as shown in the second column of Table 5, its regression coefficient is -0.013, which is still significant at the 1% level. This result further confirms that the application of big data technology can reduce the supply concentration of enterprises. Based on the above analysis, this paper assumes that H2 is confirmed.

According to the data in the third column of Table 5, without the introduction of any control variables, concentration degree of the supply chain significantly negatively affects the firm performance, and the regression coefficient is -0.079, which is significant at the level of 1%. Then, after adding many control variables, as shown in the fourth column of Table 5, its regression coefficient is -0.130, which is still significant at the 1% level. The results show that the higher the concentration of supply chain, the worse the enterprise performance. Based on the above analysis, this paper assumes that H3 is confirmed.

Table 4. Benchmark regression results

	(1) FP	(2) FP
BD	0.004^{***}	0.010^{***}
	(2.843)	(8.706)
Size		-0.004^{***}
		(-5.710)
FirmAge		-0.006^{**}
		(-2.223)
Lev		-0.151^{***}
		(-23.949)
ROA		0.830^{***}
		(49.366)
Invest		0.031^{***}
		(2.605)
Liquid		0.009^{***}
		(17.083)
SOE		-0.033^{***}
		(-19.210)
Industry & Year FE	YES	YES
_cons	0.280^{***}	0.409^{***}
	(262.197)	(24.087)
N	38228	38228
adj. R^2	0.085	0.336

t statistics in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

In the model presented in Table 5, the regression analysis following the inclusion of the interaction term in model (6) reveals a regression coefficient of -0.001 for the interaction term. This suggests that reducing the level of concentration of supply chain will strengthen the influence of big data technology application on firm performance. A lower level of concentration of supply chain leads to a more significant positive influence on the economic performance of enterprises through the implementation of digital technology. Consequently, hypothesis H4 is supported.

In Table 5, we introduce the interaction item between big data technology and the age of senior executives (BD*TMTAge) on the basis of column 3. The regression analysis results show that the coefficient of this interaction item is negative at the significance level of 1%, which means that with the increase of the average age of the senior management team, the promotion effect of big data technology on enterprise performance is weaker. The younger the average age of the senior management team, the stronger the role of big data technology in promoting corporate performance. To summarize, hypothesis H5 is supported.

Table 5. Mechanism regression results

	(1) SCC	(2) SCC	(3) FP	(4) FP	(5) FP	(6) FP
BD	-0.020***	-0.013***	0.002	0.008***	0.048***	0.076***
	(-16.541)	(-11.665)	(1.617)	(7.181)	(2.864)	(5.341)
Size		-0.034***		-0.009***		-0.004***
		(-41.565)		(-11.466)		(-4.590)
FirmAge		-0.011***		-0.007***		-0.004*
		(-3.865)		(-2.804)		(-1.744)
Lev		0.034***		-0.146***		-0.153***
		(5.109)		(-23.419)		(-24.234)
ROA		-0.141***		0.812***		0.832***
		(-9.829)		(48.320)		(49.506)
Invest		0.149***		0.050***		0.029**
		(11.112)		(4.190)		(2.413)
Liquid		0.008***		0.010***		0.009***
		(15.177)		(19.108)		(16.910)
SOE		0.006***		-0.032***		-0.031***
		(3.119)		(-18.899)		(-17.797)
SCC			-0.079***	-0.130***		
			(-13.860)	(-25.660)		
inter					-0.001***	-0.001***
					(-2.704)	(-4.722)
TMTAge					-0.004***	-0.000
					(-10.980)	(-1.252)
Industry & Year FE	YES	YES	YES	YES	YES	YES
_cons	0.311***	1.064***	0.305***	0.548***	0.454***	0.409***
	(310.140)	(56.623)	(152.520)	(31.196)	(28.631)	(20.321)
N	38228	38228	38228	38228	38228	38228
adj. R^2	0.148	0.218	0.090	0.349	0.090	0.337

4.4 Robust Analysis

In order to enhance the persuasiveness of the research results, we conducted four robustness tests, as shown in Table 6.

Initially, the dependent variable is replaced. Firm Performance_sub (FP-SUB) is used instead of the original net profit, and the results are consistent with the current study.

Second, we perform a lag test. Since the digital transformation of the core explanatory variable of this paper is obtained by capturing relevant keywords from the annual report, the total number of word frequencies is taken as the proxy variable, and big data technology is a major strategic decision for enterprises, enterprises need to evaluate their own environment and the opportunities and risks they face in order to confirm their future investment and development strategies. Therefore, in order to prevent enterprises from actually using big data technology to forecast the future layout of big data technology in the annual report, this paper deals with the core explanatory variable (BD) one stage behind and tests its relationship with FP performance level. As shown in Table 4.8 (2), after one or two periods of lag, big data technology still positively promotes enterprise performance, and it is significant above 1%, assuming H1 is still valid.

Third, the popularity of big data technology has only begun to rise rapidly in recent years, and its application fields are constantly expanding and its effectiveness is constantly improving. However, in the initial stage of the application of digital technology, it has not been widely adopted, and the application of technology is still in the process of maturity. Therefore, the sample data from 2015 to 2015 were excluded and the regression analysis was re-performed, and the results obtained were consistent with the previous results.

4.5 Heterogeneity Analysis

In the above research, this paper examines the comprehensive influence of ordinary listed companies and establishes a consistent positive correlation between them. It is important to acknowledge that significant variations exist in the attributes within the sample, potentially leading to heterogeneity in the responses of enterprises with diverse characteristics when confronted with digital technologies. Therefore, in the empirical analysis in Table 5, this research divides the samples into three groups and conducts identification tests: "traditional enterprises, platform enterprises and Internet enterprises", "high-tech enterprises - non-high-tech enterprises" and "industrial enterprises and service enterprises". The results show that the use of big data technology by traditional enterprises, platform enterprises and Internet enterprises has a significant positive impact on corporate performance. The use of big data technology by industrial and service enterprises also has a significant positive impact on business performance. In contrast, in the group of high-tech enterprises and non-high-tech enterprises, the use of big data by non-high-tech enterprises has no significant impact on enterprise performance. This paper believes that big data technology itself is high-tech, and non-high-tech enterprises may not standardize data collection and management, resulting in data missing, errors or inconsistencies. The quality of big data analysis is highly dependent on data quality, and there are problems with basic data, which will affect the effect of decision-making. Second, non-high-tech enterprises may lack professional talents and skills, big data technology needs professional data analysts to operate and interpret, and the lack of talents will lead to the technical potential cannot be released. In addition, there are application scenario factors and traditional enterprise application scenarios that may not match the core needs of the business, resulting in improper use of big data technology.

Table 6. Robust analysis results

	(1) FP_sub	(2) FP	(3) FP	(4) FP
BD	0.001			0.010^{***}
	(1.063)			(8.489)
Size	0.004^{***}	-0.005^{***}	-0.005^{***}	-0.002^{***}
	(6.840)	(-5.656)	(-5.717)	(-2.846)
FirmAge	-0.003	-0.008^{***}	-0.011^{***}	-0.008^{**}
	(-1.444)	(-2.812)	(-3.518)	(-2.444)
Lev	-0.013^{**}	-0.161^{***}	-0.162^{***}	-0.169^{***}
	(-2.138)	(-23.178)	(-21.920)	(-22.924)
ROA	2.271^{***}	0.834^{***}	0.838^{***}	0.769^{***}
	(107.316)	(46.387)	(44.182)	(40.845)
Invest	-0.035^{***}	0.040^{***}	0.045^{***}	0.043^{***}
	(-3.957)	(2.928)	(3.001)	(2.837)
Liquid	0.006^{***}	0.008^{***}	0.008^{***}	0.009^{***}
	(12.945)	(12.863)	(10.082)	(13.878)
SOE	0.009^{***}	-0.034^{***}	-0.036^{***}	-0.036^{***}
	(6.553)	(-18.776)	(-18.677)	(-18.333)
L.BD		0.010^{***}		
		(8.120)		
L2.BD			0.011^{***}	
			(7.821)	
Industry & Year FE	YES	YES	YES	YES
_cons	-0.128^{***}	0.434^{***}	0.454^{***}	0.391^{***}
	(-8.986)	(22.997)	(22.237)	(19.652)
N	38228	32115	27735	29884
adj. R^2	0.673	0.330	0.326	0.319

t statistics in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 7. Heterogeneity analysis results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	FP	FP	FP	FP	FP	FP	FP
BD	0.008***	0.013***	0.018***	-0.001	0.015***	-0.011***	0.012***
	(6.116)	(6.692)	(10.803)	(-0.451)	(4.810)	(-3.297)	(9.650)
Size	-0.003**	-0.008***	-0.009***	0.002*	-0.012***	-0.002	-0.004***
	(-2.465)	(-8.328)	(-10.380)	(1.906)	(-6.016)	(-0.911)	(-4.486)
FirmAge	-0.005	-0.008**	-0.008**	-0.000	0.035***	-0.014	-0.015***
	(-1.536)	(-2.040)	(-2.536)	(-0.054)	(5.183)	(-1.618)	(-5.180)
Lev	-0.175***	-0.102***	-0.114***	-0.204***	-0.148***	-0.138***	-0.156***
	(-20.352)	(-11.061)	(-14.893)	(-19.408)	(-8.945)	(-7.925)	(-21.209)
ROA	0.737***	1.104***	0.882***	0.750***	0.938***	1.097***	0.757***
	(36.715)	(37.229)	(41.598)	(29.123)	(21.190)	(22.537)	(38.984)
Invest	0.026*	0.050***	0.031**	0.015	-0.031	0.064**	0.036**
	(1.753)	(2.636)	(2.138)	(0.787)	(-1.149)	(2.119)	(2.453)
Liquid	0.009***	0.006***	0.005***	0.011***	0.009***	0.004**	0.010***
	(15.159)	(5.363)	(6.660)	(14.665)	(6.589)	(2.511)	(16.322)
SOE	0.000	0.000	-0.018***	-0.055***	-0.054***	-0.060***	-0.023***
	(.)	(.)	(-8.973)	(-18.975)	(-12.474)	(-13.257)	(-11.102)
Industry & Year FE	YES	YES	YES	YES	YES	YES	YES
_cons	0.388***	0.441***	0.495***	0.301***	0.465***	0.417***	0.422***
	(16.209)	(17.577)	(23.729)	(10.944)	(11.159)	(8.134)	(21.198)
N	24137	14090	22188	16040	6220	5246	26754
adj. R^2	0.299	0.353	0.327	0.336	0.351	0.347	0.347

t statistics in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

5 Conclusions and Recommendations

5.1 Research Conclusion

In the contemporary digital era, computer technologies have become the cornerstone of business innovation and transformation. The continuous miniaturization of electronic components and the exponential growth of data storage capacities have paved the way for handling massive datasets with ease. Furthermore, the advent of artificial intelligence and the Internet of Things has interconnected devices and systems on an unprecedented scale, generating a wealth of real-time data. These technological advancements have not only reshaped the business landscape but also created new opportunities for firms to leverage data-driven strategies. Against this backdrop, based on the sample data set of 1200 listed companies from 2013 to 2022, this paper delves into the intricate relationship between computer technologies and firm performance.

In the realm of computer technology, the rapid evolution of data processing capabilities, fueled by advancements in cloud computing and distributed systems, has enabled firms to harness vast amounts of data efficiently. Moreover, the development of sophisticated data mining and machine learning algorithms has empowered businesses to extract valuable insights from complex datasets, laying the groundwork for leveraging big data technologies. In this study, we employed Python for keyword frequency statistics to precisely quantify the adoption and utilization of big data technologies within these listed companies, thereby capturing the extent to which these technologies are integrated into their operations.

This study takes net profit as the index to evaluate the economic performance of enterprises, takes big data technology (as measured through our Python-based keyword frequency analysis) as the explanatory variable, and uses the ordinary least squares (OLS) linear regression model to investigate the impact of big data technology on economic performance of enterprises. The results show that the use of big data technology has a significant beneficial impact on the economic benefits of enterprises.

In addition, supply chain concentration plays a mediating role in the impact of big data technology on firm economic performance. Specifically, big data technology can reduce the concentration of the supply chain and positively promote the economic effect of enterprises. This suggests that by leveraging big data, firms are able to optimize their supply chain operations, leading to improved economic outcomes.

At the same time, the age of the senior management team plays a moderating role in the promotion of enterprise performance by big data technology. Specifically, the average age of the senior management team negatively moderates the promotion of enterprise performance by big data technology. The higher the average age, the smaller the positive impact of big data technology on enterprise performance. This highlights the importance of having a management team that is receptive and adaptable to new technologies in order to fully realize the benefits of big data.

5.2 Research Limitations and Future Prospects

Although this paper has made some achievements in studying the impact of big data technology on enterprise performance, there are still some limitations. For example, this paper does not make a classification analysis of different types of enterprises, and it is also found in the heterogeneity analysis that the research results for non-high-tech enterprises may not be fully applicable. In addition, the measurement of the application of big data technology in this paper mainly relies on text analysis, which may have a certain subjectivity. Future research can conduct targeted and in-depth research on a certain kind of enterprises. Considering the availability of data, this study takes listed companies from 2013 to 2022 as research samples, which does have certain limitations. Listed companies are different from smes in financial transparency, governance structure, financing ability and other aspects. Listed companies are usually larger and financially healthy enterprises in the industry, while smes may face higher bankruptcy risk or fiercer competition. Data of listed companies may overestimate the level of technology adoption and underestimate the impact of resource constraints. Limited to listed companies, the research conclusions can still provide important references for corporate governance and policy formulation, but the scope of application needs to be clear. In the future, it is hoped that non-listed companies can be used as research samples to conduct research, so as to draw more credible research results compared with this study. In addition, in the research method, field research and case analysis can also be combined to deeply explore the application effect of big data technology in different types of enterprises. At the same time, with the continuous development of big data technology, future research can focus on its comprehensive impact on enterprise innovation, organizational change and social responsibility, so as to provide more comprehensive theoretical support for the sustainable development of enterprises and society.

References

[1] H. W. Wen, C. Y. Wen, C. C, Lee, Impact of Digitalization and Environmental Regulation on Total Factor Productivity, *Information Economics and Policy*, Vol. 61, Article No. 101007, December, 2022. https://doi.org/10.1016/j.infoecopol.2022.101007

[2] R. K. Verma, S. Singh, A Hybrid Framework of Resource Allocation using Firefly and Deep Learning in Big Data Scheduling, *International Journal of Performability Engineering*, Vol. 20, No. 6, pp. 333-343, June, 2024. https://doi.org/10.23940/ijpe.24.06.p1.333343

[3] J. Ma, K. R. M. Raob, Efficient Resource Managing and Job Scheduling in a Heterogeneous Kubernetes Cluster for Big Data, *International Journal of Performability Engineering*, Vol. 20, No. 3, pp. 157-166, March, 2024. https://doi.org/10.23940/ijpe.24.03.p4.157166

[4] H. Ding, Z. Weng, Z. Jiang, J. Zhang, Construction of Urban Governance Portrait Based on Space-Air-Ground Integrated Digital Intelligence Pedestal, *2024 IEEE 24th International Conference on Software Quality, Reliability, and Security Companion*, Cambridge, United Kingdom, 2024, pp. 1179-1187. https://doi.org/10.1109/QRS-C63300.2024.00155

[5] Z. Wang, L. Li, B. Wang, R. Luo, L. Ye, The Application of the Teaching Model of "Promoting Learning Through Competition, Promoting Teaching Through Competition" in the Course of Data Analysis and Mining, *2024 IEEE 24th International Conference on Software Quality, Reliability, and Security Companion*, Cambridge, United Kingdom, 2024, pp. 1106-1112. https://doi.org/10.1109/QRS-C63300.2024.00146

[6] Y. Kittichotsatsawat, V. Jangkrajarng, K. Y. Tippayawong, Enhancing Coffee Supply Chain towards Sustainable Growth with Big Data and Modern Agricultural Technologies, *Sustainability*, Vol. 13, No. 8, Article No. 4593, April, 2021. https://doi.org/10.3390/su13084593

[7] R. Liu, L. H. Zheng, Z. A. Chen, M. Y. Cheng, Y. Z. Ren, Digitalization Through Supply Chains: Evidence from the Customer Concentration of Chinese Listed Companies, *Economic Modelling*, Vol. 134, Article No. 106688, May, 2024. https://doi.org/10.1016/j.econmod.2024.106688

[8] J. Bughin, M. Chui, J. Manyika, Clouds, big data, and smart assets: Ten tech-enabled business trends to watch, *McKinsey Quarterly*, Vol. 56, No. 1, pp. 75-86, 2010.

[9] G. L. Li, Y. F. Shao, How Do Top Management Team Characteristics Affect Digital Orientation? Exploring the Internal Driving Forces of Firm Digitalization, *Technology in Society*, Vol. 74, Article No. 102293, August, 2023.

[10] L. Omar, D. L. F. David, F. V. Simon, P. Javier, Big Data Analytics Capabilities, Direct and Mediating Relationships with Innovative and Business Performance, *Journal of Management Analytics*, Vol. 11, No. 2, pp. 182-201, April, 2024. https://doi.org/10.1016/j.techsoc.2023.102293

[11] T. H. Davenport, J. G. Harris, *Competing on Analytics: The New Science of Winning*, Harvard business review press, 2007.

[12] A. McAfee, E. Brynjolfsson, Big Data: The Management Revolution, *Harvard Business Review*, Vol. 90, No. 10, pp. 60-68, October, 2012.

[13] X. Ge, J. Jackson, The Big Data Application Strategy for Cost Reduction in Automotive Industry, *SAE International Journal of Commercial Vehicles*, Vol. 7, No. 2, pp. 588-598, September, 2014. https://doi.org/10.4271/2014-01-2410

[14] S. F. Wamba, S. Akter, A. Edwards, G. Chopin, D. Gnanzou, How 'Big Data' Can Make Big Impact: Findings from a Systematic Review and a Longitudinal Case Study, *International Journal of Production Economics*, Vol. 165, pp. 234-246, July, 2015. https://doi.org/10.1016/j.ijpe.2014.12.031

[15] G. W. Xu, D. A. Ali, A. Bhaumik, M. S. Yang, N. Wang, The Influence of Digital Transformation in Enterprises on the Dynamics of Supply Chain Concentration: An Empirical Analysis of Chinese A-share Listed Companies, *Journal of Organizations, Technology and Entrepreneurship*, Vol. 1, No. 2, pp. 88-97, December, 2023. https://doi.org/10.56578/jote010202

[16] J. Leveling, M. Edelbrock, B. Otto, Big Data Analytics for Supply Chain Management, *2014 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, Selangor, Malaysia, 2014, pp. 918-922. https://doi.org/10.1109/IEEM.2014.7058772

[17] J. J. Liu, Research on Supply Chain Optimization Strategy of Clothing Retail Industry under the Background of Big Data, *8th International Conference on Management, Education and Information (MEICI)*, Shenyang, the People's Republic of China, 2018, pp. 56-60. https://doi.org/10.2991/meici-18.2018.12

[18] T. H. Blank, Developing Incremental Innovation in the High-tech Industry: the Effect of Age and Tenure in Research and Development Teams, *Cross Cultural & Strategic Management*, Vol. 31, No. 2, pp. 289-308, April, 2024. https://doi.org/10.1108/CCSM-04-2023-0054

[19] C. A. O'Reilly III, M. Tushman, *Lead and Disrupt: How to Solve the Innovator's Dilemma*, Stanford University Press, 2021.

[20] H. Zhang, X. H. Wang, M. W. Akhtar, Digital Transformation, Supplier Concentration, and CEO Financial Experience: Unveiling the Dynamics of Innovation Performance in Chinese Firms, *Journal of Cleaner Production*, Vol. 442, No. 140825, February, 2024. https://doi.org/10.1016/j.jclepro.2024.140825

Biographies

Ying Dong, born in 1982, holds a doctoral degree, with a research focus on enterprise digital transformation. She has published several papers in CSSCI.

Sun Wei, born in 1982, holds a master degree, with a research focus on information engineering management and human resource management. He is currently a senior engineer.