Enterprise digital development and capacity utilization

China Accounting and Finance Review

459

Received 21 April 2023
Revised 29 February 2024
Accepted 21 March 2024

Hui Ma
Tongji University, Shanghai, China
Shenglan Chen
Zhejiang University of Technology, Hangzhou, China
Xiaoling Liu
Xiamen University, Xiamen, China, and
Pengcheng Wang
Zhejiang University of Technology, Hangzhou, China

Abstract

Purpose – To enrich the research on the economic consequences of enterprise digital development from the perspective of capacity utilization.

Design/methodology/approach – Using a sample of listed firms from 2010 to 2020, this paper exploits text analysis of annual reports to construct a proxy for enterprise digital development.

Findings – Results show that enterprise digital development not only improves their own capacity utilization but also generates a positive spillover effect on the capacity utilization of peer firms and firms in the supply chain. Next, based on the incomplete information about market demand and potential competitors when making capacity-building decisions, the mechanism tests show that improving the accuracy of market forecasts and reducing investment surges are potential channels behind the baseline results. Cross-sectional tests show the baseline result is more pronounced when industries are highly homogeneous and when firms have access to less information.

Originality/value – This paper contributes to the research related to the economic consequences of digital development. With the development of the digital economy, the real effects of enterprise digital development have also triggered extensive interest and exploration. Existing studies mainly examine the impact on physical operations, such as specialization division of labor, innovation activities, business performance or total factor productivity (Huang, Yu, & Zhang, 2019; Yuan, Xiao, Geng, & Sheng, 2021; Wang, Kuang, & Shao, 2017; Li, Liu, & Shao, 2021; Zhao, Wang, & Li, 2021). These studies measure the economic benefits from the perspective of the supply (output) side but neglect the importance of the supply system to adapt to the actual market demand. In contrast, this paper focuses on capacity utilization, aimed at estimating the net economic effect of digital development by considering the supply-demand fit scenario. Thus, our findings enrich the relevant studies on the potential consequences of digital development.

Keywords Digital development, Capacity utilization, Market demand, Investment surge, Product markets

Paper type Research paper

1. Introduction

The scale of China's digital economy accounts for 38.6% of GDP in 2020, and that percentage is expected to exceed 50% by 2025, with the era of the digital economy in full swing. During the "14th Five-Year Plan" period, cloud computing, big data and other digital industries are

Professor Hui Ma acknowledges financial support from the National Natural Science Foundation of China (No. 72102166), the MOE Project of Key Research Institute of Humanities and Social Science in University (22JJD790094) and 111 Project (B18033).

China Accounting and Finance Review
Vol. 26 No. 4, 2024
pp. 459-484
Emerald Publishing Limited
e-ISSN: 2307-3055
p-ISSN: 1029-807X
DOI 10.1108/CAFR-04-2023-0046

expected to drive more than 60 tri yuan of economic output. However, overcapacity has been a serious problem for past economic growth. The main cause of overcapacity is that we focus purely on supply output without considering its suitability for buyer demand, consequently causing the supply growth rate to far exceed the consumption growth rate. In this sense, estimating the net economic effect of digitalization requires taking into account the appropriateness of supply and demand. Therefore, this study examines the effect of enterprise digital development on capacity utilization, shedding important light on the economic logic of digital development contributing to high-quality development.

It is noteworthy that the improvement of capacity utilization relies on the optimal allocation of resources, which in turn depends on more available information to reduce uncertainty. To this end, enterprises need to enhance their ability to identify and capture key information about market preferences, demand shifts and purchasing power in advance. Prior studies show that forming a reasonable forecast of the project outlook can avoid the detachment of completed capacity investment from market demand and low capacity utilization (Paraskevopoulos, Karakitsos, & Rustem, 1991; Wang *et al.*, 2017). On the other hand, enterprises have to grasp the dynamic information of competitors to optimize their own resource allocation (Paraskevopoulos *et al.*, 1991; Wang *et al.*, 2017). In a competitive market, if the enterprise is only based on its own information without incorporating the information of potential rivals, the final resource allocation will deviate from the general equilibrium of its market (industry). Consequently, the total capacity of the market (industry) is much larger than the market demand, resulting in lower capacity utilization (Lin, 2007; Lin, Wu, & Xing, 2010). In this sense, given that digital development contributes to extracting more valuable information, we expect that enterprise digital development can enhance enterprise capacity utilization.

By collecting and analyzing various data resources to enhance the information set and reshape the information structure, digital technology can help improve enterprises' ability to predict market demand trends and competitive dynamics, thus optimizing decision-making judgments accordingly (Louck *et al.*, 2019). Specifically by exploiting digital technology to collect and analyze customers' historical transaction data and browsing footprints, enterprises can identify key market information such as consumer preferences, demand shifts, potential purchasing power and target customers to refine target markets. Thus, it is possible to form more accurate forecasts of future sales performance, inventory management, product innovation and product upgrades (Miklós-Thal & Tucker, 2019; Milgrom & Tadelis, 2019). Based on this reasoning, the development of digital technologies contributes to more accurate market forecasts, thus improving capacity utilization.

Moreover, digital technology helps to enhance the ability of enterprises to obtain and analyze information about potential competitors and then optimize their investment decisions. Gathering and analyzing information about potential competitors' dynamics is indispensable for future strategic planning, such as what types of new products or services may be introduced by a firm. The acquisition of such information helps to enhance the firm's perception and prediction of the overall output of the market and then, optimize its resource inputs in order to maximize its revenues. When detailed information about customers, contracts and products is available, firms have incentives to coordinate with their competitors in order to achieve greater gains in the product market (Goncharov & Peter, 2019; Bourveau, She, & Žaldokas, 2020). For example, Bernard, Blackburne, and Thornock (2020) show that firms have incentives to utilize relevant information technology tools to interpret the annual reports of peer firms' annual reports in order to obtain more valuable information to guide their own investment decisions by implementing investment differentiation strategies. According to the logic of the analysis, we predict that digital technologies provide firms with greater access to information on competitor dynamics and then, the easing of the investment boom can increase capacity utilization.

Using a sample of listed Chinese firms from 2010 to 2020, this paper examines the effect of enterprise digitization development on capacity utilization. The results show that enterprise digital development significantly improves capacity utilization. Specifically, for every one standard deviation increase in digitization degree, enterprise capacity utilization increases by about 4% on average. We acknowledge that the above findings may be confounded by endogeneity issues. For example, the fundamental characteristics are significantly different among firms with different levels of digitalization, which results in omitted-variable bias. It is also possible that firms with higher capacity utilization also have more resources to invest in digital technology that is a reverse causation issue. To strengthen the causal inference, we employ propensity score matching, instrumental variables and the exogenous shocks method to mitigate the endogeneity issue. Results are robust after addressing endogoenity issues and other robustness tests.

Next, this paper adopts a two-stage approach to analyze the underlying mechanism through which digital development improves capacity utilization – the reduction of information opacity related to market demand and rivalry dynamics. Specifically, we use management forecasts and the investment surge phenomenon as mechanism variables and find that enterprise digital development can significantly improve management forecast accuracy and reduce investment surge risk. Then, the improvement of management forecasting accuracy and the reduction of investment surge risk lead to an increase in capacity utilization. The step-by-step mechanism test supports the basic logic of our baseline result. Additionally, cross-sectional tests related to industry and firm characteristics show that when the industry is highly homogeneous, supported by industrial policy and executives have poorer access to information, the positive effect of digital development on capacity utilization is more pronounced.

As Gu, Sanders, and Venkateswaran (2017) state, the decisions of any firm are bound to have an impact on other firms in the network of economic linkages, suggesting that the positive effect of digital development on capacity utilization might not be limited to the firm itself. Competitive relations and supply chain relationships are two of the most important ways in which firms' behavior can spill over (Dye, 1990; Bushee & Leuz, 2005; Shroff, 2017). Accordingly, we further analyze whether enterprise digital development has spillover effects on peer firms and firms in the supply chain. Results show that focal firms' digital development can also significantly increase the capacity utilization of peer firms and firms in the supply chain. These results support the positive externalities of digital development.

This paper contributes to several lines of research. First, this paper contributes to the research related to the economic consequences of digital development. With the development of the digital economy, the real effects of enterprise digital development have also triggered extensive interest and exploration. Existing studies mainly examine the impact on physical operations, such as specialization division of labor, innovation activities, business performance or total factor productivity (Huang *et al.*, 2019; Yuan *et al.*, 2021; Wang *et al.*, 2017; Li *et al.*, 2021; Zhao *et al.*, 2021). These studies measure the economic benefits from the perspective of the supply (output) side but neglect the importance of the supply system to adapt to the actual market demand. In contrast, this paper focuses on capacity utilization, aimed at estimating the net economic effect of digital development by considering the supply-demand fit scenario. Thus, our findings enrich the relevant studies on the potential consequences of digital development.

One contemporary study is Zhao and Ren (2023), which also examines the effect of enterprise digital development on capacity utilization in a sample of Chinese firms. Our study differs from theirs in the following two ways: First, whereas their study examines the effect of enterprise digital development, it does not speak to the spillover effect of a focal firm's digital development on peer firms and firms along the supply chain, an important argument and finding in our study. Second, in their analysis, innovation is a potential mechanism through which digital development increases capacity utilization, but our study examines the channels of information transparency related to market demand and rivalry dynamics.

Second, this paper enriches the research on capacity utilization. Taking incomplete information as a logical starting point for analysis, demand uncertainty and investment surges due to a lack of information about potential entrants are important drivers of low capacity utilization (Malmendier & Tate, 2005; Xu & Zhou, 2015; Lin *et al.*, 2010). Current studies examining the improvement of capacity utilization examine the impact of institutional changes such as simplifying administrative approvals and setting up administrative approval centers (Wu & Liu, 2018; Bian & Bai, 2021; Li, Luo, & Pang, 2020), but pay little attention to the improvement of the information environment. Based on the context of the era of digital economy development, this paper explores and identifies a novel governance mechanism to enhance capacity utilization – digital development. This provides more evidence to understand how the information environment affects capacity utilization.

Third, the findings of this paper are of great practical significance for how to enhance the digital economy to empower high-quality development. Clarifying the impact of digital development on capacity utilization and its mechanisms is conducive to summarizing the experience of digital technology, empowering real decision-making and optimizing the decision-making of resource allocation. Therefore, our findings are helpful to provide policy insights on how to strengthen scientific and technological development and accelerate digital transformation, thus consequently achieving high-quality development in the economy.

2. Literature review and research hypotheses

2.1 Literature review

How to improve capacity utilization and avoid overcapacity is an important issue for economic development, and it has received close attention from academics. As its definition suggests, as the ratio of market-digested capacity to potential capacity, increasing capacity utilization necessarily involves the main participants of economic activities: enterprises, markets, governments and so on at the same time (Stiglitz, 1999; Xu & Zhou, 2015).

The decisive role of the market essentially reflects the guiding role of market operation information in resource allocation. However, the existence of information friction in the real world impairs the quality of decision-making, such as through low capacity utilization. The difficulty in grasping complete information on market demand ex ante causes firms to overestimate market demand and form more optimistic market forecasts. As such, the construction of capacity is greater than that required under equilibrium conditions and ultimately, resulting in relatively low capacity utilization (Paraskevopoulos *et al.*, 1991; Malmendier & Tate, 2005; Lu & Poddar, 2006). Consistent with this argument, a theoretical study by Goyal and Netessine (2007) shows that firms' capacity investment decisions are made before the market demand uncertainty is resolved, and firms tend to make larger capacity investments. Xu and Zhou (2015) find that even if the industry's development prospects are relatively certain, uncertain market demand makes high-efficiency firms relatively cautious about capacity investment, which in turn leaves more room for inefficient firms to invest in capacity, ultimately resulting in low overall capacity utilization. From a market demand perspective, Xiao and Zheng (2018) also find that improving supply quality to meet the needs of upgrading the demand structure is an effective way to resolve overcapacity.

In addition, information about potential competitors is also an important factor that cannot be ignored in the firm's decision function (Beatty, Liao, & Yu, 2013; Bernard *et al.*, 2020). Bernard *et al.* (2020) found that firms actively search and study the annual reports of their peers to extract valuable information and then use it to guide their own investment decisions, which ultimately leads to significant improvements in both their own and their peers' investment efficiency. More importantly, the firm will adopt a differentiated product strategy to avoid severe competition. Bourveau *et al.* (2020) show that when firms disclose more detailed information about their customers and products in financial reports, firms in the same industry coordinate each other's behaviors in the product market in order to maintain their respective higher profit margins. On the contrary, the lack of relevant information about potential competitors may cause firms' investment decisions to exhibit a tidal wave phenomenon, leading to ex-post low capacity utilization (Lin, 2007; Lin *et al.*, 2010). In the analytical framework of investment surge, when a large number of firms form a high degree of consensus on the development prospects and expected profits, the unknown information on the decision-making behavior of potential competitors triggers the phenomenon of investment surge, which makes the market's (the industry) aggregated capacity is greater the market demand and ultimately results in low capacity utilization.

Prior studies also show that government is also a factor explaining low capacity utilization in China. This is because local governments have strong incentives to intervene in corporate investment under the system of fiscal decentralization and the political promotion of officials centered on the assessment of GDP growth. Additionally, local governments have the ability to subsidize corporate investment, thus causing a surge in corporate investment and triggering low capacity utilization (Jiang, Geng, Lv, & Li, 2012). Xu and Ma (2019) show that under the incentive of political promotion, local officials will urge firms to increase the overcapacity in their jurisdictions and consequently cause a decline in capacity utilization when they are approaching a position change. It is worth noting that the essential reason is the lack of information on the behaviors of competitor firms. Thus, it is difficult to coordinate competitive strategies with each other.

Relevant studies have also actively explored specific mechanisms that can help improve the capacity utilization of firms. However, most prior studies focus on institutional reforms such as administrative approval reforms, regional market integration, the Belt and Road Initiative and mixed ownership (Wu & Liu, 2018; Bian & Bai, 2021; Li *et al.*, 2020) and pay little attention to effective measures adopted by firms themselves. Different from existing studies, based on the digital economy era, we identify and explore market-based mechanisms to prevent the risk of overcapacity from the perspective of digital development, which provides a new method to enhance capacity utilization.

2.2 Hypothesis development

Facing the rapid development of the digital economy, firms are increasingly adopting digital technologies such as big data, cloud computing, the Internet of Things, social media and mobile platforms. The economic rationale behind digital development is to improve the transparency of the information environment by improving the efficiency of data collection, utilization and information processing flow. If so, it is possible for firms to enhance their ability to perceive ahead of time, thus optimizing decision-making. That is, digital development is the process of extracting information, creating knowledge and realizing the value of data resources (Ma, Jin, & Wang, 2021). In this sense, the expansion and even reshaping of the information set available to the firm is an essential benefit for digital development, which stems from both the expansion of the quantity and the improvement of the quality of the data (Rozario & Zhang, 2021).

The acquisition and analysis of more available data is the basis for firms to predict market demand and changes in competitive dynamics in advance. Digital technology promotes information exchange and learning between firms and their suppliers, customers, governments and competitors and thus, opens up the flow of information in the economic links of research and development, production, circulation, distribution and consumption, realizing all-channel, all-chain and all-process information connectivity. For firms, more available information acquired through digital development is helpful to fill in the blind spots of data and bridge the data gap to form a complete data chain. Not surprisingly, firms can obtain more information output for market demand forecasting, competitor analysis and supply chain management, which in turn serves their own project screening and evaluation, production and sales and other operational decisions (Liu, Chen, & Chou, 2011; Constantiou & Kallinikos, 2015; Meng & Wang, 2020).

Tucker, Foldesy, Roos, and Rodt (2020) point out that digital technology is an effective way to cope with the current fast-changing business environment and artificial intelligence, machine learning and advanced algorithms can help firms reveal unanticipated factors behind changes in performance, such as industry tidal waves and supply chain risks, enabling firms to make better responses. Ding, Lev, Peng, Sun, and Vasarhelyi (2020) found that the accuracy of accounting estimates for insurance claim-related items is significantly higher when using machine learning methods. Similarly, Babina, Fedyk, He, and Hodson (2021) confirmed that the use of digital technology can help firms better capture customers' latent demand preferences and consumption types, thus accelerating product development, enriching product portfolios and ultimately, realizing higher sales performance. Bernard *et al.* (2020) show that when firms make their own investment decisions, they have incentives to utilize information technology tools to actively capture and extract information about their competitors, thereby increasing investment efficiency and performance.

Meanwhile, digital development is helpful to improve the quality of information. Specifically, the adoption of digital technology means greater automation and intelligence when information is recorded, collected and exchanged between firms and internal and external parties, thus helping to identify data anomalies and reduce errors (Seow, Goh, Pan, Yong, & Chek, 2021). Tucker *et al.* (2020) point out that digital tools can enhance the quality of financial reporting by minimizing human error, cognitive bias and increasing accounting consistency and enforcing relevant standards in the preparation of financial reports. Moreover, digital technology contributes to greater immediacy, transparency and verifiability of data (Dai & Vasarhelyi, 2017; Gaur & Gaiha, 2020). For example, unlike traditional accounting bookkeeping, blockchain provides a decentralized distributed ledger database, which provides real-time records and verifications of the flow of funds and information about the current transactions of each node, etc. and the real-time nature leaves few opportunities to manipulate financial statements (Chiu, 2021). More importantly, once the information is verified and added to the blockchain, it is stored permanently and cannot be modified, which means that the data relying on the blockchain has higher stability and reliability (Korpela, Hallikas, & Dahlberg, 2017; Chod, Trichakis, Tsoukalas, Aspegren, & Weber, 2020).

Based on the above analysis, digital development helps firms increase the availability and quality of information related to market demand and competitors, thus optimizing their capacity investment decisions and improving capacity utilization. Accordingly, we propose the following research hypotheses:

H1. Other things being equal, corporate digital development can improve capacity utilization.

3. Research design and data

3.1 Sample selection and data sources

We use the data of Chinese A-share listed firms from 2010 to 2020 as the initial research sample and then, execute the following screening procedures: (1) deleting the samples of firms in the financial and insurance industries; (2) deleting the samples of firms in the year of listing and (3) deleting the samples with missing data. After the above screening, our final sample includes 26,966 firm-year observations. Financial data are from the CSMAR database and provincial-level data are from the Statistical Yearbook database. To mitigate the effect of outliers, all continuous variables are winsorized at the top and bottom percentiles.

3.2 Key variables and regression model

3.2.1 Capacity utilization. The capacity utilization of a firm refers to the ratio of its actual output to its production capacity. Specifically, we adopt the transcendental logarithmic cost function method to measure a firm's actual and optimal outputs and the ratio of the actual output to the optimal output is the capacity utilization (*CU*). The greater the value of the *CU*, the better the utilization of corporate production capacity. Using this method to estimate the capacity utilization of firms is common in recent studies (Xu & Ma, 2019; Ma *et al.*, 2021).

3.2.2 Enterprise digital development. Digitization is a systematic process that fundamentally relies on the layout and development of such core technologies as artificial intelligence, blockchain, cloud computing, big data, etc. (Qi & Xiao, 2020). Recent studies construct firm-level digitization indicators by using text analysis of corporate annual reports with manually defined digitized keywords (Wu, Hu, Lin, & Ren, 2021; Yuan *et al.*, 2021). Following prior studies, we take the corporate digitization keyword set defined by Wu *et al.* (2021) as the seed word set; then we expand the above seed word set with similar words using the Word2Vec machine learning technique, as shown in Appendix 1. Finally, we construct the enterprise digitization variables by using the digitization vocabulary word frequency and the sentence frequency, respectively. Specifically, *DigitalW* is the ratio of the total number of digitization vocabulary in the text of MD&A to the total vocabulary in the corresponding text; *DigitalS* is the ratio of the total number of sentences containing digitization vocabulary to the total number of sentences in the text of MD&A. The larger the value of *DigitalW* and *DigitalS*, the higher the degree of corporate digitization. We acknowledge that these two variables constructed by text analysis of the annual report have some limitations. For example, firms can strategically use more words related to digital development in their annual reports but not put resources into digital technology that is strategic disclosure.

3.2.3 Regression model. To examine the impact of enterprise digitization on capacity utilization, we estimate the following regression model:

$$CU_{i,t} = \alpha_0 + \beta_1 Digital_{i,t-1} + Controls_{i,t-1} + Fixed\,Effects + \varepsilon \tag{1}$$

where *CU* is the capacity utilization variable and *Digital* is the enterprise digitization variables *DigitalW* and *DigitalS*. According to existing studies (Xu & Ma, 2019; Ma *et al.*, 2021; Fang & Mao, 2021), the control variables *Controls* include firm size (*Size*), leverage (*Leverage*), return on assets (*ROA*), cash flow from operations (*CFO*), total factor productivity (*TFP*), number of employees (*Employee*), firm age (*Age*), board size (*BSize*), independent director ratio (*IndR*), shareholding ratio of the first largest shareholder (*SH1*), nature of ownership (*SOE*) and industry concentration (*HHI*). To mitigate the endogeneity problem, the key explanatory and control variables were lagged by one period. We also include firm-fixed effects and year-fixed effects. The regression coefficient β_1 measures the impact of enterprise digitization development on future capacity utilization and we expect β_1 to be significantly positive. For a detailed definition of variables, please see Appendix 2.

3.3 Descriptive statistics

Table 1 reports the descriptive statistics of the main variables used in this paper. The mean of corporate capacity utilization (*CU)* is 0.785, the median is 0.682 and the standard deviation is 0.145, which is similar to the descriptive results of the existing studies (Xu & Ma, 2019). The mean values of enterprise digitization development variables *DigitalW* and *DigitalS* are 0.004 and 0.060, respectively. The mean of *Size* is 22.149, the mean of *Leverage* is 0.436, the

Variable	N	Mean	SD	Q1	Median	Q3
CU	26,966	0.785	0.145	0.479	0.682	0.962
DigitalW	26,966	0.004	0.031	0.000	0.001	0.004
DigitalS	26,966	0.060	0.249	0.006	0.025	0.072
Size	26,966	22.149	1.292	21.230	21.988	22.893
Leverage	26,966	0.436	0.213	0.267	0.427	0.591
ROA	26,966	0.046	0.077	0.014	0.040	0.077
CFO	26,966	0.044	0.071	0.006	0.044	0.086
TFP	26,966	14.544	1.735	13.161	14.343	15.734
Employee	26,966	7.642	1.273	6.802	7.591	8.440
Age	26,966	2.918	0.327	2.708	2.944	3.135
BSize	26,966	2.289	0.251	2.197	2.303	2.485
IndR	26,966	0.381	0.072	0.333	0.364	0.429
SH1	26,966	0.342	0.148	0.226	0.320	0.443
SOE	26,966	0.363	0.481	0.000	0.000	1.000
HHI	26,966	0.047	0.051	0.020	0.037	0.059

Note(s): This table reports summary statistics for the variables used in our baseline regression. *CU,* or Capacity utilization is defined as the ratio of the actual output to the optimal output. *DigitalW* is the ratio of the total number of digitization vocabulary in the text of MD&A to the total vocabulary of MD&A. *DigitalS* is the ratio of the total number of sentences containing digitization vocabulary to the total number of sentences in the text of MD&A. *Size* is the natural logarithm of total assets. *Leverage* is the ratio of book value of total debt to total assets. *ROA* is the return on assets, defined as the ratio of earnings to total assets. *CFO* is the ratio of cash flow from operations to total assets. *TFP* is the total factor productivity estimated by the LP approach. *Employee* is the natural logarithm of the total number of employees. *Age* is the natural log of the number of years since a firm went public. *BSize* is the natural logarithm of the total number of independent directors. *IndR* is the ratio of independent directors to total directors. *SH1* is the ratio of shares held by the largest shareholder to the total number of shares. *SOE* equals one if a firm is an SOE and zero otherwise. *HHI* is a Herfindahl index variable based on firms' sales revenues
Source(s): Authors' own work

Table 1.
Summary statistics

mean of *ROA* is 0.046, the mean of *CFO* is 0.044, the mean of *TFP* is 14.544 and the average shareholding of the largest shareholder is 34.2%. The means of *SOE* and *HHI* were 0.363 and 0.047, respectively.

4. Empirical tests and results

4.1 Baseline: enterprise digital development and capacity utilization

We first examine how enterprise digital development affects its capacity utilization. Results are reported in Table 2. Columns (1) and (2) report the results of enterprise digital development measured by the percentage of digitalized word frequency and we find that coefficients of *DigitalW* are positive and significant, indicating that the enterprise digital development increases its capacity utilization. The result is also economically significant. For example, in column (2), a one-standard deviation increase in *DigitalW* leads to a 4.95% increase in capacity utilization (CU). When using the sentence frequency with digitalized keywords as a measure of digitization level, the coefficients on *DigitalS* in columns (3) and (4) remain significantly positive. Collectively, these results in Table 4 show that, both in terms of statistical significance and economic significance, enterprise digital development has a significant improvement effect on capacity utilization, which is consistent with our hypothesis.

4.2 Addressing endogeneity issue

The above results suggest that firms' digital development can significantly improve their capacity utilization, but this may be disturbed by endogeneity issues. For example, the

	CU			
	(1)	(2)	(3)	(4)
DigitalW	*2.722****	*1.850***		
	(3.09)	*(2.30)*		
DigitalS			*0.183****	*0.150****
			(3.35)	*(2.61)*
Size		−0.078***		−0.078***
		(−6.77)		(−6.79)
Leverage		0.196***		0.196***
		(5.96)		(5.96)
ROA		0.792***		0.792***
		(19.09)		(19.08)
CFO		0.411***		0.412***
		(10.51)		(10.52)
TFP		0.079***		0.079***
		(13.77)		(13.75)
Employee		−0.002		−0.002
		(−0.20)		(−0.21)
Age		0.211***		0.211***
		(4.07)		(4.05)
BSize		0.017		0.017
		(1.60)		(1.58)
IndR		0.015		0.015
		(0.54)		(0.55)
SH1		−0.001		−0.001
		(−0.01)		(−0.02)
SOE		−0.000		0.000
		(−0.00)		(0.01)
HHI		−0.163***		−0.162***
		(−3.12)		(−3.12)
Constant	0.775***	0.587**	0.774***	0.596**
	(239.01)	(2.34)	(236.69)	(2.37)
Firm FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
N	26,966	26,966	26,966	26,966
Adj. R^2	0.745	0.772	0.745	0.772

Note(s): This table presents the regression results of capacity utilization on enterprise digital development. The dependent variable is capacity utilization *(CU)*. The key independent variables are enterprise digital development proxies, *DigitalW* and *DigitalS*. Control variables include *Size, Leverage, ROA, CFO, TFP, Employee, Age, BSize, IndR, SH1, SOE and HHI.* Regressions in columns 1 and 3 exclude firm-level controls and regressions in columns 2 and 4 include firm-level controls. Variable definitions are presented in Appendix 2. Continuous variables are winsorized at the 1 and 99% levels. Standard errors are all heteroscedastic, robust and clustered at the firm level. *, ** and *** indicate statistical significance at the 10, 5 and 1% levels, respectively, using a two-tailed *t*-test
Source(s): Authors' own work

Table 2.
Digital development and capacity utilization

fundamental characteristics (e.g. financial status, information environment, competitive condition and executive capabilities) may be significantly different among firms with different digital developments, and these characteristics may affect capacity utilization, thus resulting in omitted variable bias. Additionally, firms with high capacity utilization may also have stronger incentives and financial resources to invest in digital technology, leading to reverse causality. To strengthen the causal inference, we use the following three approaches to address the endogeneity problem:

4.2.1 Propensity score matching (PSM). Whether a firm invests in digitization technology and digital development is endogenous to firm characteristics rather than randomized. This means that firms with different levels of digital development may also differ significantly in fundamentals, thus resulting in different capacity utilization. To alleviate this concern, we first use PSM to balance fundamentals across firms with different levels of digital development. Specifically, we define firms with digital development in the upper third quartile as a treated sample ($Treat = 1$) and remaining firms as a control sample ($Treat = 0$). Next, we run the Logit model using the dummy variable for digital development level ($Treat$) as the explanatory variable and the control variables in the regression model (1) as the explanatory variables. Finally, based on the resulting propensity scores, we define the firms with a low digital development level that most closely match firms with a high digital development level as the control group.

Panel A in Table 3 provides the descriptive statistics after PSM, and we find that the firm characteristics do not differ significantly between the treated group and the control group. Next, we re-examine the baseline using the PSM-matched sample and the results of Panel B show that the coefficients of *DigitalW* and *DigitaS* are significantly positive. This implies that the main inference holds when using the PSM sample to address the endogeneity issue.

4.2.2 Instrumental variable approach. Moreover, we use the instrumental variable approach to mitigate the endogeneity problem. Following Huang *et al.* (2019) and Yuan *et al.* (2021), we use each city's postal and telecommunication development in 1984 as an instrumental variable for enterprise digitization development. A city's communication development process will affect the application and acceptance of information technology by the enterprise, satisfying the relevance condition. Moreover, the development of post and telecommunications is a social infrastructure for the public and does not have a direct effect on the capacity utilization of enterprises, which means that the instrumental variable can also be exogenous.

However, it is noteworthy that each city's postal and telecommunication development in 1984 is cross-sectional data and cannot be directly used as an instrumental variable of the panel data. Specifically, our instrumental variable (phone) is the interaction term of the number of fixed-line telephones per 10,000 people in each prefectural level in 1984 and the annual number of Internet users in the whole country, which is consistent with prior studies (Zhao *et al.*, 2021; Yuan *et al.*, 2021). The results of the instrumental variable test are shown in Table 4. Panel A shows that the regression coefficients of instrumental variables are significantly positive at the 1% level and the Cragg–Donald Wald F-statistic is greater than 10. This suggests that the instrumental variable is highly correlated with the endogenous explanatory variable, and the null hypothesis of a weak instrumental variable is rejected. Panel B reports the results of second-stage regression, showing that the coefficients of *DigitalW* and *DigitalS* are still significantly positive.

Overall, Table 4 indicates that the positive relationship between enterprise digitalization development and capacity utilization still holds after using instrumental variables to mitigate endogeneity issues.

4.2.3 Other robustness tests. In addition, we perform the following tests to ensure the robustness of the findings: (1) following, we construct a C-D production function to estimate the capacity utilization; (2) use the natural logarithm of the vocabulary frequency of the firm's digitization keywords or the frequency of the sentences that contain the keywords to measure the firm's digitization development level; (3) use standardized digitization development variables and (4) to control for the interference of relevant characteristics at the provincial level or industry level, provincial×year and provincial×year×industry fixed effects in the model. The results of the robustness tests reported in Table 5 are similar to Table 4, i.e. the digitization development contributes to capacity utilization.

Panel A: Summary statistics after matching

Variables	HDigital (N = 8972) Mean	Median	LDigital (N = 8972) Mean	Median	Mean diff	Median diff
Size	22.211	22.042	22.213	22.051	−0.002	−0.008
Leverage	0.421	0.415	0.422	0.410	−0.001	0.005
ROA	0.052	0.046	0.053	0.043	−0.001	0.003
CFO	0.046	0.047	0.046	0.044	0.001	0.002
TFP	14.590	14.410	14.591	14.372	−0.001	0.039
Employee	7.788	7.739	7.792	7.712	−0.004	0.027
Age	2.904	2.944	2.905	2.944	0.000	0.000
BSize	2.285	2.303	2.282	2.303	0.003	0.000
IndR	0.384	0.364	0.384	0.364	0.000	0.000
SH1	0.342	0.321	0.343	0.319	−0.001	0.002
SOE	0.325	0.000	0.327	0.000	−0.002	0.000
HHI	0.047	0.037	0.047	0.037	0.000	0.000

Panel B: Regression result

	CU (1)	CU (2)
DigitalW	*1.669**	
	(1.83)	
DigitalS		*0.125***
		(2.24)
Size	−0.094***	−0.094***
	(−5.83)	(−5.86)
Leverage	0.261***	0.261***
	(6.91)	(6.90)
ROA	0.786***	0.786***
	(14.12)	(14.12)
CFO	0.462***	0.462***
	(9.04)	(9.05)
TFP	0.083***	0.083***
	(10.98)	(10.96)
Employee	−0.002	−0.002
	(−0.15)	(−0.16)
Age	0.185***	0.184***
	(2.95)	(2.92)
BSize	0.012	0.011
	(0.80)	(0.78)
IndR	0.012	0.012
	(0.31)	(0.32)
SH1	0.018	0.018
	(0.27)	(0.27)
SOE	0.010	0.010
	(0.36)	(0.37)
HHI	−0.176***	−0.176***
	(−2.69)	(−2.68)
Constant	0.997***	1.009***
	(3.21)	(3.25)
Firm FEs	Yes	Yes
Year FEs	Yes	Yes
N	17,944	17,944
Adj. R^2	0.790	0.790

Note(s): This table presents the regression results of capacity utilization on enterprise digital development using a PSM sample. Panel A reports the summary statistics of the main variables after matching. Panel B presents the regression results of capacity utilization on enterprise digital development using a PSM sample. The dependent variable is capacity utilization *(CU)*. The key independent variables are enterprise digital development proxies, *DigitalW* and *DigitalS*. Control variables include *Size, Leverage, ROA, CFO, TFP, Employee, Age, BSize, IndR, SH1, SOE and HHI.* Variable definitions are presented in Appendix 2. Continuous variables are winsorized at the 1and 99% levels. Standard errors are all heteroscedastic, robust and clustered at the firm level. *, ** and *** indicate statistical significance at the 10, 5 and 1% levels, respectively, using a two-tailed t-test
Source(s): Authors' own work

Table 3.
Propensity score match

Panel A: First-stage results		
Dep Var =	(1) *DigitalW*	(2) *DigitalS*
Phone	*0.001****	*0.008****
	(4.83)	*(5.16)*
F-statistics	23.35	26.66

Panel B: Second-stage results	CU	
	(1)	(2)
DigitalW	*29.394***	
	(1.96)	
DigitalS		*1.693***
		(1.99)
Size	−0.197***	−0.199***
	(−20.77)	(−22.02)
Leverage	0.367***	0.367***
	(10.99)	(11.08)
ROA	0.903***	0.910***
	(12.88)	(13.59)
CFO	0.426***	0.431***
	(7.23)	(7.24)
TFP	0.344***	0.345***
	(28.19)	(28.72)
Employee	−0.059***	−0.058***
	(−6.20)	(−6.31)
Age	0.006	0.004
	(0.33)	(0.22)
BSize	0.010	0.006
	(0.70)	(0.43)
IndR	−0.045	−0.041
	(−0.91)	(−0.85)
SH1	0.127***	0.123***
	(3.16)	(3.17)
SOE	0.064***	0.066***
	(4.07)	(4.07)
HHI	0.094	0.095
	(1.46)	(1.49)
Firm FEs	Yes	Yes
Year FEs	Yes	Yes
N	26,966	26,966
Adj. R^2	0.257	0.270

Note(s): This table presents the instrumental variable regression results of capacity utilization on enterprise digital development. Panel A presents the first-stage regressions in which the dependent variable is one of the enterprise digital development proxies, *DigitalW* and *DigitalS*. The instrumental variable, *phone*, is the interaction term of the number of fixed-line telephones per 10,000 people in each prefectural level in 1984 and the annual number of Internet users in the whole country. Panel B reports the second-stage regressions in which the dependent variable is capacity utilization *(CU)* and the key dependent variable is the fitted value of the enterprise digital development measure. Control variables include *Size, Leverage, ROA, CFO, TFP, Employee, Age, BSize, IndR, SH1, SOE and HHI.* Regressions in columns 1 and 3 exclude firm-level controls and regressions in columns 2 and 4 include firm-level controls. Variable definitions are presented in Appendix 2. Continuous variables are winsorized at the 1 and 99% levels. Standard errors are all heteroscedastic, robust and clustered at the firm level. *, ** and *** indicate statistical significance at the 10, 5 and 1% levels, respectively, using a two-tailed *t*-test
Source(s): Authors' own work

Table 4.
Instrumental variable approach

	CU (1)	(2)
Panel A: Capacity utilization based on C-D product function		
DigitalW	*1.948***	
	(2.20)	
DigitalS		*0.137***
		(2.53)
Controls	Yes	Yes
Firm FEs	Yes	Yes
Year FEs	Yes	Yes
N	26,966	26,966
Adj. R^2	0.741	0.741
Panel B: The logarithm of digital development		
DigitalW	*0.005**	
	(1.88)	
DigitalS		*0.006**
		(1.95)
Controls	Yes	Yes
Firm FEs	Yes	Yes
Year FEs	Yes	Yes
N	26,966	26,966
Adj. R^2	0.745	0.772
Panel C: Standardized digital development		
DigitalW	*0.003**	
	(1.94)	
DigitalS		*0.003***
		(1.96)
Controls	Yes	Yes
Firm FEs	Yes	Yes
Year FEs	Yes	Yes
N	26,966	26,966
Adj. R^2	0.772	0.772

Panel D: High-dimension fixed effects

Dep Var =	CU (1)	(2)	(3)	(4)
DigitalW	*1.416**		*1.393**	
	(1.74)		*(1.93)*	
DigitalS		*0.107***		*0.112***
		(2.11)		*(2.43)*
Controls	Yes	Yes	Yes	Yes
Year×Province	Yes	Yes	–	–
Year×Province×Industry	–	–	Yes	Yes
Firm FEs	Yes	Yes	Yes	Yes
N	27,609	27,609	27,597	27,597
Adj. R^2	0.770	0.770	0.792	0.792

Note(s): This table reports the robustness check of the baseline. Panel A reports the regression results of constructing a C-D production function to estimate capacity utilization. Panel B reports the regression results of using the natural logarithm of the frequency of vocabulary the firm's digitization keywords or the frequency of the sentences that contain the keywords to measure the firm's digitization development level. Panel C reports the regression results of using standardized digitization development variables. Panel D reports regression results of adding high-dimension fixed effects. The dependent variable is capacity utilization *(CU).* The key independent variables are enterprise digital development proxies, *DigitalW* and *DigitalS*. Control variables include *Size, Leverage, ROA, CFO, TFP, Employee, Age, BSize, IndR, SH1, SOE and HHI.* Variable definitions are presented in Appendix 2. Continuous variables are winsorized at the 1 and 99% levels. Standard errors are all heteroscedastic, robust and clustered at the firm level. *, ** and *** indicate statistical significance at the 10, 5 and 1% levels, respectively, using a two-tailed *t*-test
Source(s): Authors' own work

Table 5.
Other robustness check

5. Mechanism tests

We further conducted tests to explore the mechanism through which digital development affects capacity utilization. A low capacity utilization can be intuitively interpreted as a firm's actual output being less than the optimal capacity, i.e. a part of the firm's production capacity has not been used. The reason for this is that incomplete mastery of key information such as market demand and preference changes can lead to the formation of less accurate market forecasts, thus resulting in a mismatch between supply and demand and supply-side overcapacity (Paraskevopoulos *et al.*, 1991; Xiao & Zheng, 2018). Put differently, inaccurate forecasts of market demand induce lower capacity utilization. In addition, it is difficult for enterprises to anticipate and judge reasonable investment due to incomplete information about the capacity decisions of peer firms in the industry and the resulting investment surge and overcapacity accordingly (Lin, 2007; Lin *et al.*, 2010). Zhao, Wang, and Liu (2017) found that even in the same region, enterprises can hardly overcome this investment surge risk due to incomplete information. Based on the above analysis, we analyze the mechanism through which enterprise digital development enhance capacity utilization from the perspective of market forecast accuracy and investment surge.

Regarding the research design of the mechanism test, we adopt the two-stage approach of Di Giuli and Laux (2021). In the first stage, we examine the effect of enterprise digitization on the mechanism variable (M). In the second stage, we examine how the predicted mechanism variable ($\widehat{M}$) derived in the first stage affects capacity utilization. The two-stage approach actually disaggregates the total effect in Table 4 into two parts. In simple terms, the two-stage test can be realized by estimating the following system of equations:

$$M = \alpha_0 + \beta_1 Digital + Controls + Fixed\,Effects + \varepsilon \tag{2}$$

$$CU = \alpha_0 + \gamma_1 \widehat{M} + Controls + Fixed\,Effects + \varepsilon \tag{3}$$

where M is the mechanism variable, *Digital* is the enterprise digitization development, *CU* is the capacity utilization variable and *Controls* is the control variable consistent with the control variables used in model (1). Equation (2) is the first stage test, β_1 indicates the effect of enterprise digitization on the mechanism variable and equation (3) is the second stage test, γ_1 indicates the effect of the mechanism variable driven by the digitization development on the enterprise's capacity utilization.

5.1 Enterprise digitization and accuracy of market forecasts

The management forecast of future market prospects is an important factor affecting resource allocation. Once management makes an incorrect estimate of the return generated by an investment project, it will also distort its investment decision-making distortions: over-optimism about future market demand tends to overestimate investment returns and thus lead to over-investment (Goodman, Neamtiu, Shroff, & White, 2014). To explore this mechanism, we use the accuracy of management forecasts as a proxy for market forecast accuracy. The enterprise does not provide forecast data on sales that directly reflects market demand and management forecasts have taken market demand, project profitability and other factors into account. Specifically, *forecast* takes the value of 1 if the management forecast is close to changes in EBITDA to proxy the forecast accuracy and otherwise it is 0. If enterprise digitalization can help the enterprise better identify market demand and enhance the accuracy of market prediction, it can be expected that β_1 in the model (5) will be significantly positive and γ_1 in the model (6) will be significantly positive.

Table 6 provides the regression results. The regression results in columns (1) and (2) show that the coefficients of *DigitalW* and *DigitalS* in the first-stage regression are significantly positive, suggesting that the accuracy of management forecasts increases with enterprise

	The first-step(Forecast)		The second-step (CU)	
	(1)	(2)	(3)	(4)
DigitalW	*1.920**			
	(1.78)			
DigitalS		*0.143***		
		(2.12)		
$\widehat{ForecastW}$			*1.086****	
			(2.71)	
$\widehat{ForecastS}$				*0.100****
				(3.04)
Size	−0.030***	−0.031***	−0.045***	−0.048***
	(−2.95)	(−2.97)	(−2.73)	(−3.17)
Leverage	0.192***	0.192***	−0.012	0.005
	(5.94)	(5.94)	(−0.14)	(0.07)
ROA	0.270***	0.272***	0.504***	0.527***
	(4.84)	(4.86)	(4.39)	(5.48)
CFO	0.175***	0.176***	0.226***	0.242***
	(3.38)	(3.39)	(2.82)	(3.49)
TFP	−0.005	−0.005	0.084***	0.083***
	(−0.96)	(−0.98)	(13.66)	(14.00)
Employee	−0.008	−0.008	0.006	0.005
	(−0.91)	(−0.94)	(0.57)	(0.51)
Age	0.092	0.091	0.108*	0.117*
	(1.54)	(1.52)	(1.70)	(1.93)
BSize	0.039**	0.038**	−0.023	−0.020
	(2.35)	(2.34)	(−1.25)	(−1.23)
IndR	0.022	0.022	−0.009	−0.007
	(0.43)	(0.45)	(−0.31)	(−0.25)
SH1	−0.223***	−0.222***	0.244**	0.224**
	(−4.09)	(−4.09)	(2.23)	(2.35)
SOE	0.010	0.010	−0.012	−0.011
	(0.40)	(0.41)	(−0.61)	(−0.57)
HHI	0.060	0.061	−0.226***	−0.221***
	(0.84)	(0.85)	(−3.87)	(−3.88)
Constant	0.690***	0.700***	−0.158	−0.100
	(2.67)	(2.71)	(−0.43)	(−0.30)
Firm FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
N	26,966	26,966	26,966	26,966
Adj. R^2	0.111	0.111	0.770	0.770

Note(s): This table reports two-stage regression analyses on whether enterprise digital development promotes management forecast accuracy and in turn, increases capacity utilization. Columns (1) and (2) report the first-stage regression estimates for the effect of digital development on forecast accuracy. The dependent variable is forecast accuracy, Forecast, which takes the value of 1 if the management forecast is close to changes in EBITDA to proxy the forecast accuracy and otherwise it is 0. The key independent variables are digitization development proxies, *DigitalW* and *DigitalS*. Control variables include *Size, Leverage, ROA, CFO, TFP, Employee, Age, BSize, IndR, SH1, SOE and HHI.* Columns (3) and (4) report the second-stage regression estimates for the effect of digtial-related forecast accuracy on capacity utilization. The dependent variable is capacity utilization, *CU*. The key independent variables, $\widehat{ForecastW}$ ($\widehat{ForecastS}$) are the predicted *DigitalW* *(DigitalS)* based on the first-stage regression, capturing the part of forecast accuracy that can be explained by digital development. Control variables are the same as those in the first-stage regression. For brevity, we do not report the coefficients of the control variables. Standard errors are all heteroscedastic, robust, clustered at the firm level. Appendix 2 presents variable definitions. Continuous variables are winsorized at the 1 and 99% levels. *, ** and *** represent significance levels of 10, 5 and 1%, respectively
Source(s): Authors' own work

Table 6.
Mechanism test: Forecast accuracy

digital development, i.e. that digital development helps firms better identify market demand. Meanwhile, the second-stage regression results show that the coefficients on $\widehat{ForecastW}(\widehat{ForecastS})$ are both significantly positive, which indicates that forecasts with higher accuracy can significantly improve capacity utilization. Taken together, the results of Table 6 indicate that enterprise digital development achieves higher capacity utilization by improving the accuracy of market demand forecasts.

5.2 Firm digitization levels and investment surges

We adopt a similar approach to examine whether investment surge mitigation is an important mechanism through which enterprise digital development enhance capacity utilization. First, we follow the methodology of Jing and Zhang (2021) to calculate the investment similarity *InvSimi* between focal firms and peer firms in the same industry to measure the extent of investment surges that firms may experience. The investment surge refers to the fact that a large number of identical or similar investments have been made among firms. There is a high degree of investment similarity. If enterprise digital development helps to improve the investment surge caused by enterprises' incomplete information about the total investment in the industry, then it can be expected that β_1 in model (2) will be significantly negative. Meanwhile, γ_1 in model (3) should be significantly negative.

The results of the two-stage regression are given in Table 7. Columns (1) and (2) first provide the results of the tests, with the investment surge variable (*InvSimi*) as the explanatory variable. We find that in the first stage, the coefficients on *DigitalW* and *DigitalS* are significantly negative at the 1% level, which indicates that enterprise digital development lowers their investment similarity. In columns (3) and (4), the coefficients on $\widehat{InvSimi}W(\widehat{InvSimi}S)$ are significantly negative in the second stage, suggesting that the reduction of investment similarity driven by enterprise digitalization development significantly enhances capacity utilization. The mechanism test results in Table 7 confirm that investment surge mitigation is an important mechanism through which enterprise digital development improves capacity utilization.

6. Cross-sectional tests

In this section, we conduct cross-sectional tests based on industry and firm characteristics to further support the above inference.

First, we focus on the influence of industry homogeneity on cross-sectional variation in the impact of the digital development level on capacity utilization. Industry homogeneity refers to a high degree of consistency in terms of technological capabilities, target customers, product mix and resources for firms in the same industry environment (Cairney & Young, 2006). Mauri and Michaels (1998) find that firms in industries with more homogeneity tend to be highly homogeneous in terms of investing in R&D, marketing and capital investment. As such, for firms in the industry with high homogeneity, managers are more likely to form a common perception of market demand forecasts and potential investment opportunities. Thus, if enterprise digital development improves capacity utilization by incorporating more information on market demand and competitor actions, it can be expected that the improvement will be more obvious in firms with higher homogeneity.

To this end, we follow Parrino's (1997) methodology to use the mean value of the correlation coefficient between stock returns and industry returns for each firm within the same industry to measure the industry homogeneity variable (Homo), and then distinguish industries into two groups of high and low industry homogeneity based on the sample median value. As shown in Table 8, we find that for enterprises in industries with high

	The first-step(Inv_Simi)		The second-step (CU)	
	(1)	(2)	(3)	(4)
DigitalW	*−0.180****			
	(−3.64)			
DigitalS		*−0.012****		
		(−3.82)		
$\widehat{InvSimi}W$			*−0.103***	
			(−2.30)	
$\widehat{InvSimi}S$				*−0.112****
				(−2.61)
Size	0.000	0.000	−0.075***	−0.074***
	(0.89)	(0.91)	(−6.40)	(−6.39)
Leverage	0.012***	0.012***	0.325***	0.336***
	(9.99)	(9.96)	(5.04)	(5.42)
ROA	0.021***	0.021***	1.010***	1.029***
	(7.53)	(7.54)	(9.61)	(10.11)
CFO	0.017***	0.017***	0.590***	0.605***
	(5.78)	(5.74)	(6.81)	(7.23)
TFP	−0.003***	−0.003***	0.046***	0.043***
	(−10.99)	(−10.98)	(3.00)	(2.89)
Employee	0.000	0.000	−0.001	−0.001
	(0.40)	(0.42)	(−0.07)	(−0.06)
Age	−0.004***	−0.004***	0.174***	0.170***
	(−4.01)	(−4.00)	(3.20)	(3.11)
BSize	−0.002**	−0.002**	−0.001	−0.003
	(−2.36)	(−2.33)	(−0.10)	(−0.22)
IndR	0.003	0.003	0.041	0.044
	(1.04)	(1.04)	(1.36)	(1.45)
SH1	0.002	0.002	0.019	0.021
	(1.16)	(1.16)	(0.33)	(0.36)
SOE	−0.003***	−0.003***	−0.032	−0.035
	(−5.95)	(−6.00)	(−1.37)	(−1.51)
HHI	0.012**	0.012**	−0.042	−0.030
	(2.49)	(2.49)	(−0.55)	(−0.42)
Constant	0.091***	0.091***	1.522***	1.610***
	(12.81)	(12.79)	(3.15)	(3.39)
Firm FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
N	26,966	26,966	26,966	26,966
Adj. R^2	0.199	0.199	0.772	0.772

Note(s): This table reports two-stage regression analyses on whether enterprise digital development reduces the surge of investment and in turn, increases capacity utilization. Columns (1) and (2) report the first-stage regression estimates for the effect of digital development on the surge of investment. The dependent variable is investment similarity, *Inv_Simi*, proposed by Jing and Zhang (2021). The key independent variables are digitization development proxies, *DigitalW* and *DigitalS*. Control variables include *Size, Leverage, ROA, CFO, TFP, Employee, Age, BSize, IndR, SH1, SOE and HHI*. Columns (3) and (4) report the second-stage regression estimates for the effect of digtial-related reductions in investment surges on capacity utilization. The dependent variable is capacity utilization, *CU*. The key independent variables, $Inv\_\widehat{Simi}W$ ($Inv\_\widehat{Simi}S$) are the predicted *DigitalW* (*DigitalS*) based on the first-stage regression, capturing the part of investment similarity that can be explained by digital development. Control variables are the same as those in the first-stage regression. For brevity, we do not report the coefficients of the control variables. Standard errors are all heteroscedastic, robust, clustered at the firm level. Appendix 2 presents variable definitions. Continuous variables are winsorized at the 1 and 99% levels. *, ** and *** represent significance levels of 10, 5 and 1%, respectively
Source(s): Authors' own work

Table 7.
Mechanism test: The surge of investment

	CU			
	(1)High	(2)Low	(3)High	(4)Low
DigitalW	*3.585***	*1.074*		
	(2.00)	*(1.31)*		
DigitalS			*0.291****	*0.057*
			(2.60)	*(1.14)*
Size	−0.079***	−0.085***	−0.079***	−0.085***
	(−4.48)	(−5.76)	(−4.52)	(−5.76)
Leverage	0.086	0.254***	0.087	0.254***
	(1.58)	(6.44)	(1.60)	(6.44)
ROA	0.812***	0.738***	0.812***	0.739***
	(12.67)	(13.36)	(12.67)	(13.36)
CFO	0.509***	0.283***	0.509***	0.283***
	(9.38)	(5.52)	(9.39)	(5.53)
TFP	0.072***	0.086***	0.072***	0.086***
	(8.86)	(11.41)	(8.81)	(11.40)
Employee	−0.032**	0.005	−0.032**	0.005
	(−2.04)	(0.40)	(−2.04)	(0.39)
Age	0.071	0.320***	0.071	0.320***
	(0.60)	(5.13)	(0.60)	(5.12)
BSize	0.017	0.017	0.017	0.017
	(1.12)	(1.30)	(1.12)	(1.29)
IndR	0.028	0.007	0.028	0.007
	(0.61)	(0.21)	(0.60)	(0.22)
SH1	−0.001	0.041	−0.001	0.040
	(−0.01)	(0.58)	(−0.01)	(0.56)
SOE	−0.022	0.051**	−0.021	0.051**
	(−0.69)	(1.98)	(−0.66)	(1.99)
HHI	−0.100	−0.151**	−0.100	−0.152**
	(−1.46)	(−2.05)	(−1.45)	(−2.06)
Constant	1.409***	0.221	1.425***	0.223
	(2.98)	(0.68)	(3.01)	(0.69)
Firm FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
N	12,054	14,584	12,054	14,584
Adj. R^2	0.790	0.813	0.790	0.813

Note(s): This table examines whether the relationship between enterprise digital development and capacity utilization varies with industry homogeneity. We distinguish industries into two groups of high and low industry homogeneity based on the sample median value. The dependent variable is capacity utilization *(CU)*. The key independent variables are enterprise digital development proxies, *DigitalW* and *DigitalS*. Control variables include *Size, Leverage, ROA, CFO, TFP, Employee, Age, BSize, IndR, SH1, SOE and HHI*. Variable definitions are presented in Appendix 2. Continuous variables are winsorized at the 1% and 99% levels. Standard errors are all heteroscedastic, robust and clustered at the firm level. *, ** and *** indicate statistical significance at the 10, 5 and 1% levels, respectively, using a two-tailed *t*-test
Source(s): Authors' own work

Table 8.
Cross-sectional tests: Industry homogeneity

homogeneity, the coefficients on *DigitalW* and *DigitalS* are all significantly positive, but in the group with low industry homogeneity, the coefficients on *DigitalW* and *DigitalS* are insignificant. These results indicate that the positive effect of digital development on capacity utilization mainly exists in enterprises that have a high level of industry homogeneity.

Second, we explore the impact of firms' information acquisition capabilities. If the positive effect of digital development on capacity utilization is due to increased available information on market demand and potential rivals, then we expect that firms with poorer information acquisition capabilities will have a more pronounced baseline result. We use the centrality of the alumni relationship networks of firms' executives and directors as a measure of their information acquisition capabilities because a large body of sociological research suggests that when an individual is central in network relationships, he/she has an advantage in information acquisition and processing (Newman, 2010). Faleye, Kovacs, and Venkateswaran (2014) found that CEOs who are centrally located in their alumni relationship networks have a more pronounced information advantage and face relatively lower uncertainty, thus making more R&D investments and producing high-quality patented products. Javakhadze, Ferris, and French (2016) provide evidence that CEOs that are more central to their networks can make investment efficiency higher.

Specifically, following prior studies, we first calculate the four network centrality variables of corporate executives (Freeman, 1977): Degree, Closeness, Betweenness and Eigenvector. Then, we conduct a principal component analysis (PCA) with the four centrality indicators mentioned above and use the first major principal component as a measure to construct the corporate executive centrality variable. Finally, we divide the samples into firms with higher centrality and those with low centrality based on the industry median value. The results of Table 9 show that they are in line with the prediction.

7. Spillover effects of firms' digital development

Gu *et al.* (2017) state that the decisions of any one firm will inevitably have an impact on other firms in the network of economic linkages. The spillover effects of corporate disclosure and business decisions among peer firms have been investigated in theoretical and empirical studies (Dye, 1990; Bushee & Leuz, 2005; Shroff, 2017). Thus, the impact of digital development may not be limited to the focal firms themselves but may spill over to other firms as well. We examine the spillover effects of enterprise digital development on peers and firms in the supply chain. Specifically, the improved market forecasts and investment decisions due to enterprises' digital development can help peers and firms in the supply chain make more accurate estimations of market demand and production capacity, thus increase capacity utilization.

In order to test the spillover effect, we construct *Peer_ DigitalW* (*Peer_ DigitalS*), the average digitization degree of peer firms. Columns (1) to (2) of Table 10 report the regression results, and the coefficients on *DigitalW and DigitalS* are both significantly positive. That is, enterprise digital development improves its own capacity utilization. More importantly, the coefficients on *Peer_ DigitalW and Peer_ DigitalS* are also significantly positive, suggesting that after controlling for enterprise digitization development, the digitalization development of peers can also significantly improve capacity utilization. Taking column (2) as an example, every one standard deviation increase in *Peer_ DigitalS* will increase enterprise capacity utilization by 7.74%, which is economically significant.

In addition, the supply chain is an important economic channel through which spillovers arise (Cohen & Frazzini, 2008). Firms need information from other firms in their supply chain to plan their operational business more efficiently. For example, Radhakrishnan, Wang, and Zhang (2014) found that high-quality information provided by client firms can help firms make better decisions on capacity inputs, production scheduling and inventory management, which ultimately enhances performance. Chiu, Lin, Tsai, and Teh (2017) found that firms' investment efficiency significantly improves when clients disclose more information about their business risks. Similarly, a large body of research on supply chains finds that one important reason for firms' overinvestment is customers' overestimated demand for their products (Lee, Padmanabhan, & Whang, 1997). Thus, we expect that enterprise digital development has a positive spillover effect on firms in the supply chain.

To test this spillover effect, we follow Ke, Li, and Zhang (2020) to use input–output table relationships to identify the supply chain industry of the firm's industry and then, calculate

	CU			
	(1)High	(2)Low	(3)High	(4)Low
DigitalW	*0.582*	*2.811****		
	(0.37)	*(2.58)*		
DigitalS			*0.077*	*0.173****
			(0.78)	*(2.60)*
Size	−0.068***	−0.115***	−0.068***	−0.115***
	(−3.66)	(−4.99)	(−3.68)	(−4.99)
Leverage	0.154**	0.241***	0.154**	0.241***
	(2.31)	(3.51)	(2.31)	(3.51)
ROA	0.758***	0.876***	0.757***	0.876***
	(10.81)	(11.86)	(10.81)	(11.85)
CFO	0.453***	0.286***	0.454***	0.286***
	(7.42)	(4.20)	(7.43)	(4.21)
TFP	0.069***	0.082***	0.069***	0.081***
	(7.14)	(8.11)	(7.14)	(8.10)
Employee	−0.021	0.008	−0.021	0.007
	(−1.07)	(0.47)	(−1.08)	(0.45)
Age	0.174**	0.388***	0.172**	0.388***
	(2.09)	(4.20)	(2.06)	(4.20)
BSize	0.004	0.016	0.004	0.016
	(0.26)	(0.91)	(0.26)	(0.90)
IndR	−0.013	0.024	−0.013	0.025
	(−0.30)	(0.49)	(−0.30)	(0.50)
SH1	−0.102	−0.007	−0.102	−0.007
	(−1.04)	(−0.07)	(−1.03)	(−0.07)
SOE	0.031	−0.009	0.031	−0.009
	(1.13)	(−0.32)	(1.14)	(−0.30)
HHI	0.009	−0.064	0.010	−0.064
	(0.12)	(−0.74)	(0.13)	(−0.74)
Constant	0.853**	0.747	0.865**	0.752
	(1.97)	(1.59)	(1.99)	(1.60)
Firm FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
N	9,733	6,914	9,733	6,914
Adj. R^2	0.815	0.818	0.815	0.818

Note(s): This table examines whether the relationship between enterprise digital development and capacity utilization varies with a firm's information acquisition. We distinguish industries into two groups of high and low information acquisition based on the sample median value. The dependent variable is capacity utilization *(CU)*. The key independent variables are enterprise digital development proxies, *DigitalW* and *DigitalS*. Control variables include *Size, Leverage, ROA, CFO, TFP, Employee, Age, BSize, IndR, SH1, SOE and HHI*. Variable definitions are presented in Appendix 2. Continuous variables are winsorized at the 1 and 99% levels. Standard errors are all heteroscedastic, robust and clustered at the firm level. *, ** and *** indicate statistical significance at the 10, 5 and 1% levels, respectively, using a two-tailed *t*-test
Source(s): Authors' own work

Table 9.
Cross-sectional tests: Information acquisition

the average digitization degree mean *SupplyChain_DigitalW* (*SupplyChain_DigitalS*) of firms in the supply chain. As shown in columns (3) and (4) of Table 10, the coefficients for *SupplyChain_DigitalW* and *SupplyChain_DigitalS* are both significantly positive after controlling for the firm's own digital development, suggesting a positive spillover effect of enterprise development in the supply chain. The results in Table 10 show that digitalization has positive externalities on the capacity utilization of other firms within the industry and in the supply chain.

	CU			
	Spillover effect on peers		Spillover effect on supply chain	
	(1)	(2)	(3)	(4)
Peer_DigitalW	*4.525***			
	(2.22)			
Peer_DigitalS		*0.244**		
		(1.89)		
Supplychain_DigitalW			*5.706***	
			(2.52)	
Supplychain_DigitalS				*0.315***
				(2.16)
DigitalW	6.112***		1.602**	
	(2.78)		(2.02)	
DigitalS		0.351**		0.105**
		(2.52)		(2.10)
Size	−0.097***	−0.097***	−0.109***	−0.109***
	(−7.37)	(−7.39)	(−8.86)	(−8.86)
Leverage	0.265***	0.265***	0.269***	0.269***
	(8.12)	(8.12)	(8.79)	(8.78)
ROA	0.808***	0.809***	0.830***	0.831***
	(18.38)	(18.39)	(19.28)	(19.28)
CFO	0.464***	0.464***	0.434***	0.434***
	(11.22)	(11.24)	(10.90)	(10.92)
TFP	0.085***	0.085***	0.121***	0.121***
	(12.88)	(12.87)	(16.67)	(16.67)
Employee	0.005	0.005	0.002	0.001
	(0.39)	(0.39)	(0.15)	(0.14)
Age	0.217***	0.217***	0.210***	0.210***
	(3.81)	(3.79)	(3.99)	(3.98)
BSize	0.018	0.018	0.018*	0.017
	(1.64)	(1.63)	(1.65)	(1.63)
IndR	0.009	0.009	0.009	0.008
	(0.31)	(0.30)	(0.31)	(0.30)
SH1	0.073	0.071	0.058	0.058
	(1.14)	(1.12)	(0.99)	(0.98)
SOE	−0.020	−0.020	−0.020	−0.019
	(−0.98)	(−0.97)	(−1.00)	(−0.99)
HHI	−0.142**	−0.142**	0.067	0.068
	(−2.47)	(−2.47)	(1.28)	(1.30)
Constant	0.767***	0.776***	0.554**	0.560**
	(2.67)	(2.69)	(2.05)	(2.07)
Firm FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
N	26,966	26,966	21,427	21,427
Adj. R^2	0.791	0.791	0.799	0.799

Note(s): This table presents the regression results of the spillover effect of capacity utilization on enterprise digital development. The dependent variable is capacity utilization *(CU)*. The key independent variables are peer firms' or supply-chain firms' digital development, *Peer_DigitalW* (*Peer_DigitalS) and Supplychain_DigitalW* (*Supplychain_DigitalS)*. Focal firms' digital development proxies are *DigitalW* and *DigitalS*. Control variables include *Size, Leverage, ROA, CFO, TFP, Employee, Age, BSize, IndR, SH1, SOE and HHI.* Variable definitions are presented in Appendix 2. Continuous variables are winsorized at the 1 and 99% levels. Standard errors are all heteroscedastic, robust and clustered at the firm level. *, ** and *** indicate statistical significance at the 10, 5 and 1% levels, respectively, using a two-tailed *t*-test
Source(s): Authors' own work

Table 10. The spillover effect on peers and firms in supply chain

8. Conclusions

This paper examines the impact of enterprise digital development on capacity utilization. Using a sample of listed enterprises from 2010 to 2020, we use text-based analysis techniques to construct enterprise digital development indicators and then, empirically examine the impact of enterprise digital development on its capacity utilization and the potential mechanism.

The results show that enterprise digital development can significantly enhance its capacity utilization. The mechanism tests reveal that the accuracy of market demand forecasts and the reduction of investment surges are important channels contributing to the increase in capacity utilization. In addition, cross-sectional tests show that the baseline result is more pronounced when industries are highly homogeneous and when firms have access to less information. Finally, this paper also finds that enterprise digital development also significantly increases the capacity utilization of peer firms and firms in the supply chain, with significant spillover effects.

This paper provides relevant empirical evidence that enterprise digital development enhances capacity utilization, which not only enriches the relevant research on the economic consequences of digital transformation but also has practical significance for the country to further promote enterprise digital development and establish a long-term mechanism to prevent overcapacity. The government should adopt more policies to support firms in developing digital technology, thus increasing capacity utilization and achieving high-quality development.

References

Babina, T., Fedyk, A., He, A., & Hodson, J. (2021). Artificial intelligence, firm growth, and product innovation. Working Paper.

Beatty, A., Liao, S., & Yu, J. J. (2013). The spillover effect of fraudulent financial reporting on peer firms' investments. *Journal of Accounting and Economics*, *55*(2-3), 183–205. doi: 10.1016/j.jacceco.2013.01.003.

Bernard, D., Blackburne, T., & Thornock, J. (2020). Information flows among rivals and corporate investment. *Journal of Financial Economics*, *136*(3), 760–779. doi: 10.1016/j.jfineco.2019.11.008.

Bian, Y., & Bai, J. (2021). Can regional market integration improve firms' capacity utilization (in Chinese). *Financial Research*, *47*(11), 64–77.

Bourveau, T., She, G., & Žaldokas, A. (2020). Corporate disclosure as a tacit coordination mechanism: Evidence from cartel enforcement regulations. *Journal of Accounting Research*, *58*(2), 295–332. doi: 10.1111/1475-679x.12301.

Bushee, B. J., & Leuz, C. (2005). Economic consequences of SEC disclosure regulation: Evidence from the OTC bulletin board. *Journal of Accounting and Economics*, *39*(2), 233–264. doi: 10.1016/j.jacceco.2004.04.002.

Cairney, T. D., & Young, G. R. (2006). Homogenous industries and auditor specialization: An indication of production economies. *Auditing: A Journal of Practice and Theory*, *25*(1), 49–67. doi: 10.2308/aud.2006.25.1.49.

Chiu, T. T. (2021). Blockchain adoption and investment sensitivity to stock price. Working Paper.

Chiu, C. K., Lin, C. P., Tsai, Y. H., & Teh, S. F. (2017). Enhancing knowledge sharing in high-tech firms: The moderating role of collectivism and power distance. *Cross Cultural and Strategic Management*, *25*(3), 468–491. doi: 10.1108/ccsm-03-2017-0034.

Chod, J., Trichakis, N., Tsoukalas, G., Aspegren, H., & Weber, M. (2020). On the financing benefits of supply chain transparency and blockchain adoption. *Management Science*, *66*(10), 4378–4396. doi: 10.1287/mnsc.2019.3434.

Cohen, L., & Frazzini, A. (2008). Economic links and predictable returns. *The Journal of Finance*, *63*(4), 1977–2011. doi: 10.1111/j.1540-6261.2008.01379.x.

Constantiou, I. D., & Kallinikos, J. (2015). New games, new rules: Big data and the changing context of strategy. *Journal of Information Technology*, *30*(1), 44–57. doi: 10.1057/jit.2014.17.

Dai, J., & Vasarhelyi, M. A. (2017). Toward blockchain-based accounting and assurance. *Journal of Information Systems*, *31*(3), 5–21. doi: 10.2308/isys-51804.

Di Giuli, A., & Laux, P. A. (2021). The effect of media-linked directors on financing and external governance. *Journal of Financial Economics*, *145*(2), 103–131. doi: 10.1016/j.jfineco.2021.07.017.

Ding, K., Lev, B., Peng, X., Sun, T., & Vasarhelyi, M. A. (2020). Machine learning improves accounting estimates: Evidence from insurance payments. *Review of Accounting Studies*, *25*(3), 1098–1134. doi: 10.1007/s11142-020-09546-9.

Dye, R. A. (1990). Mandatory versus voluntary disclosures: The cases of financial and real externalities. *Accounting Review*, *65*(1), 1–24.

Faleye, O., Kovacs, T., & Venkateswaran, A. (2014). Do better-connected CEOs innovate more?. *Journal of Financial and Quantitative Analysis*, *49*(5-6), 1201–1225. doi: 10.1017/s0022109014000714.

Fang, S., & Mao, Q. (2021). Human capital expansion and firm capacity utilization - evidence from China's 'university enrollment expansion' (in Chinese). *Economics (Quarterly)*, *21*(06), 1993–2016.

Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry*, *40*(1), 35–41. doi: 10.2307/3033543.

Gaur, V., & Gaiha, A. (2020). Building a transparent supply chain blockchain can enhance trust, efficiency, and speed. *Harvard Business Review*, *98*(3), 94–103.

Goncharov, I., & Peter, C. D. (2019). Does reporting transparency affect industry coordination? Evidence from the duration of international cartels. *The Accounting Review*, *94*(3), 149–175. doi: 10.2308/accr-52201.

Goodman, T. H., Neamtiu, M., Shroff, N., & White, H. D. (2014). Management forecast quality and capital investment decisions. *The Accounting Review*, *89*(1), 331–365. doi: 10.2308/accr-50575.

Goyal, M., & Netessine, S. (2007). Strategic technology choice and capacity investment under demand uncertainty. *Management Science*, *53*(2), 192–207. doi: 10.1287/mnsc.1060.0611.

Gu, T., Sanders, N. R., & Venkateswaran, A. (2017). CEO incentives and customer-supplier relations. *Production and Operations Management*, *26*(9), 1705–1727. doi: 10.1111/poms.12715.

Huang, Q., Yu, Y., & Zhang, S. (2019). Internet development and manufacturing productivity enhancement: Intrinsic mechanism and Chinese experience (in Chinese). *China Industrial Economy*, *377*(08), 5–23.

Javakhadze, D., Ferris, S. P., & French, D. W. (2016). Social capital, investments, and external financing. *Journal of Corporate Finance*, *37*, 38–55. doi: 10.1016/j.jcorpfin.2015.12.001.

Jiang, F., Geng, Q., Lv, D., & Li, X. P. (2012). Regional competition, institutional distortion and the formation mechanism of overcapacity (in Chinese). *China Industrial Economy*, 06, 44–56.

Jing, W., & Zhang, X. (2021). Online social networks and corporate investment similarity. *Journal of Corporate Finance*, *68*, 101921. doi: 10.1016/j.jcorpfin.2021.101921.

Ke, R., Li, M., & Zhang, Y. (2020). Directors' informational role in corporate voluntary disclosure: An analysis of directors from related industries. *Contemporary Accounting Research*, *37*(1), 392–418. doi: 10.1111/1911-3846.12522.

Korpela, K., Hallikas, J., & Dahlberg, T. (2017). Digital supply chain transformation toward blockchain integration. Working Paper.

Lee, H. L., Padmanabhan, V., & Whang, S. (1997). Information distortion in a supply chain: The bullwhip effect. *Management Science*, *43*(4), 546–558. doi: 10.1287/mnsc.43.4.546.

Li, Q., Liu, L., & Shao, J. (2021). Digital transformation, supply chain integration and firm performance-the moderating effect of entrepreneurship (in Chinese). *Economic Management*, *43*(10), 5–23.

Li, Q., Luo, F., & Pang, T. (2020). Can the belt and road initiative alleviate overcapacity in Chinese firms? (in Chinese). *Industrial Economics Research*, 04, 129–142.

Lin, Y. (2007). Tidal wave phenomenon and the reconstruction of macroeconomic theory in developing countries (in Chinese). *Economic Research*, 01, 126–131.

Lin, Y., Wu, H., & Xing, Y. (2010). The 'surge phenomenon' and the formation mechanism of overcapacity (in Chinese). *Economic Research*, *45*(10), 4–19.

Liu, D. Y., Chen, S. W., & Chou, T. C. (2011). Resource fit in digital transformation: Lessons learned from the CBC Bank global e-banking project. *Management Decision*, *49*(10), 1728–1742. doi: 10.1108/00251741111183852.

Loucks, J., Macaulay, J., Noronha, A., & Wade, M. (2019). *Digital vortex: How today's market leaders can beat disruptive competitors at their own game*. Lausanne: IMD International.

Lu, Y., & Poddar, S. (2006). The choice of capacity in mixed duopoly under demand uncertainty. *The Manchester School*, *74*(3), 266–272. doi: 10.1111/j.1467-9957.2006.00492.x.

Ma, H., Jin, Q., & Wang, X. (2021). Big data and accounting functions (in Chinese). *Journal of Management Science*, 09, 1–17.

Malmendier, U., & Tate, G. (2005). CEO overconfidence and corporate investment. *The Journal of Finance*, *60*(6), 2661–2700. doi: 10.1111/j.1540-6261.2005.00813.x.

Mauri, A. J., & Michaels, M. P. (1998). Firm and industry effects within strategic management: An empirical examination. *Strategic Management Journal*, *19*(3), 211–219. doi: 10.1002/(sici)1097-0266(199803)19:33.0.co;2-t.

Meng, F., & Wang, W. (2020). Research on the mechanism of digitalization to the improvement of manufacturing enterprises performance based on mediating effect. Working Paper.

Miklós-Thal, J., & Tucker, C. (2019). Collusion by algorithm: Does better demand prediction facilitate coordination between sellers?. *Management Science*, *65*(4), 1552–1561. doi: 10.1287/mnsc.2019.3287.

Milgrom, P. R., & Tadelis, S. (2019). How artificial intelligence and machine learning can impact market design. In *The Economics of Artificial Intelligence: An Agenda*, 567–585.

Newman, M. E. (2010). *Networks: An introduction*. Oxford: Oxford University Press.

Paraskevopoulos, D., Karakitsos, E., & Rustem, B. (1991). Robust capacity planning under uncertainty. *Management Science*, *37*(7), 787–800. doi: 10.1287/mnsc.37.7.787.

Parrino, R. (1997). CEO turnover and outside succession a cross-sectional analysis. *Journal of Financial Economics*, *46*(2), 165–197. doi: 10.1016/s0304-405x(97)00028-7.

Qi, I., & Xiao, X. (2020). Corporate management change in the era of digital economy (in Chinese). *Management World*, *36*(06), 135–152.

Radhakrishnan, S., Wang, Z., & Zhang, Y. (2014). Customers' capital market information quality and suppliers' performance. *Production and Operations Management*, *23*(10), 1690–1705. doi: 10.1111/poms.12211.

Rozario, A., & Zhang, C. A. (2021). The effects of artificial intelligence on firms' internal information auality. Working Paper.

Seow, P. S., Goh, C., Pan, G., Yong, M., & Chek, J. (2021). Embracing digital transformation in accounting and finance. Working Paper.

Shroff, N. (2017). Corporate investment and changes in GAAP. *Review of Accounting Studies*, *22*(1), 1–63. doi: 10.1007/s11142-016-9375-x.

Stiglitz, J. E. (1999). Reforming the global economic architecture: Lessons from recent crises. *The Journal of Finance*, *54*(4), 1508–1521. doi: 10.1111/0022-1082.00154.

Tucker, J., Foldesy, J., Roos, A., Rodt, M., & (2020). How digital CFOs are transforming Finance. Working Paper.

Wang, Y., Kuang, X., & Shao, W. B. (2017). Informatization, corporate flexibility and capacity utilization (in Chinese). *World Econ*, *40*(01), 67–90.

Wu, L., & Liu, C. (2018). Project matching and overcapacity in China (in Chinese). *Economic Research*, *53*(10), 67–81.

Wu, F., Hu, H., Lin, H., & Ren, X. (2021). Firms' digital transformation and capital market performance - empirical evidence from stock liquidity (in Chinese). *Management World*, *37*(07), 130–144.

Xiao, M., & Zheng, Y. (2018). Can supply quality improvement resolve China's manufacturing overcapacity? –Based on the perspective of structural optimization and technological progress (in Chinese). *China Soft Science*, 12, 126–139.

Xu, Y., & Ma, G. (2019). Local officials change and corporate overcapacity (in Chinese). *Economic Research*, *54*(05), 129–145.

Xu, C., & Zhou, N. (2015). Endogenous changes in market structure and the governance of overcapacity (in Chinese). *Economic Research*, *50*(02), 75–87.

Yuan, C., Xiao, T., Geng, C., & Sheng, Y. (2021). Digital transformation and corporate division of labor: Specialization or vertical integration (in Chinese). *China Industrial Economics*, *402*(09), 137–155.

Zhao, N., & Ren, J. (2023). Impact of enterprise digital transformation on capacity utilization: Evidence from China. *Plos One*, *18*(3), e0283249. doi: 10.1371/journal.pone.0283249.

Zhao, C., Wang, W., & Li, X. (2021). How digital transformation affects enterprise total factor productivity (in Chinese). *Finance and Trade Economics*, *42*(07), 114–129.

Zhao, N., Wang, B., & Liu, Y. (2017). City clusters, agglomeration effect and 'investment surge'-an empirical study based on 20 city clusters in China (in Chinese). *China Industrial Economy*, *11*, 81–99.

Appendix 1

Dimension	Keywords
Artificial intelligence	artificial intelligence, AI, intelligence, intelligent, automation, business intelligence, image understanding, image recognition, speech recognition identification, investment decision aid systems, intelligent data analysis, intelligent terminals, intelligent robots, robots, industrial robots robotics, industrial robotics, machine learning, deep learning, semantic search, biometrics, biometrics recognition, face recognition, voice recognition, identity recognition, autonomous driving voice recognition, identity recognition, autonomous driving, driverless driving, natural language processing
Blockchain	blockchain, digital currency, bitcoin, distributed computing, differential privacy technology, smart financial contracts
Cloud computing	cloud computing, graph computing, in-memory computing, multi-party secure computing, brain-like computing, green computing, cognitive computing, edge computing, cloud infection, mobility, cloud, informatization, On-Line Computing, IT, ICT, cloud platforms, IOT, networking, converged architecture, billions of concurrency, EB Storage, Internet of Things, information physical systems
Big data	big data, data mining, text mining, real-time data, data warehouse, data analytics, datamining, virtualization, data acquisition, data exchange, digitization, data convergence, data management, data middleware, data platform, data sharing, BI, data storage, data application, data driver, data center, data service, data analytics system, big data, data asset, visualization, data governance, big data application, data processing, data system, big data intelligence, data visualization, heterogeneous data, credit information, augmented reality, mixed reality, virtual reality
Application of digital technology	internet, internet+, mobile internet, industrial internet, mobile internet, networking platform, online and offline, intelligent management, intelligent decision-making, intelligent production, intelligent manufacturing, intelligent manufacturing, intelligent control, intelligent factories, deep integration, cross-border integration, internet healthcare, E-commerce, mobile payment, third party payment, NFC payment, smart energy, B2B, B2C, C2B, C2C, O2O, net connection, smart wear, smart agriculture, smart transport, smart healthcare, smart customer service, smart home, smart investment, smart tourism, smart environmental protection, smart grid, smart marketing, digital marketing, unmanned retail, internet finance, digital finance, Fintech, fintech, quantitative finance, open banking, digital economy

Source(s): Authors' own work

Table A1. Keywords related digital technology

Appendix 2

Table A2.
Variable definition

Variables	Definition
CU	Capacity utilization is defined as the ratio of the actual output to the optimal output
DigitalW	The ratio of the total number of digitization vocabularies in the text of MD&A to the total vocabulary of MD&A
DigitalS	The ratio of the total number of sentences containing digitization vocabulary to the total number of sentences in the text of MD&A
Size	The natural logarithm of total assets
Leverage	The ratio of book value of total debt to total assets
ROA	The return on assets, defined as the ratio of earnings to total assets
CFO	The ratio of cash flow from operations to total assets
TFP	Total factor productivity estimated using the LP approach
Employee	The natural logarithm of the total number of employees
Age	The natural log of the number of years since a firm goes public
BSize	The natural logarithm of the total number of independent directors
IndR	The ratio of independent directors to total directors
SH1	The ratio of shares held by the largest shareholder to the total number of shares
SOE	SOE equals one if a firm is an SOE, and zero otherwise
HHI	Herfindahl index variables based on firms' sales revenues

Source(s): Authors' own work

About the authors

Hui Ma is also affiliated with the Shanghai University of Finance and Economics, Shanghai, China. Hui Ma is the corresponding author and can be contacted at: mahuiacc@126.com

Xiaoling Liu is also affiliated with the Sun Yat-sen University, Guangzhou, China.