





ARTICLE



<https://doi.org/10.1057/s41599-025-04613-5>

OPEN

# Mixed ownership reform and digitalisation

Wei Tu<sup>1</sup> , Wei-Chiao Huang<sup>2</sup>, Nianzhai Ma<sup>3</sup> & Juan He<sup>4</sup> 

Employing a machine learning measure, we find that mixed ownership reform in state-owned enterprises (SOEs) significantly advances digitalisation. This effect is primarily achieved through the pay-performance sensitivities of management and corporate risk-taking; it is particularly pronounced in competitive industries and among corporate decision-makers who did not experience the Great Chinese Famine during childhood. Our study not only explores digitalisation measures with the help of cutting-edge natural language processing techniques but also expands the literature on digitalisation motivation and the impact of mixed ownership reform on business decisions. The findings have important implications for promoting digitalisation strategies in SOEs.

<sup>1</sup> School of Economics and Management, Guizhou Normal University, Guizhou, China. <sup>2</sup> Department of Economics, Western Michigan University, Kalamazoo, MI, USA. <sup>3</sup> School of Economics, Central University of Finance and Economics, Beijing, China. <sup>4</sup> The School of Accounting, Guizhou University of Finance and Economics, Guiyang, China. ✉email: [jhe@mail.gufe.edu.cn](mailto:jhe@mail.gufe.edu.cn)

## Introduction

Digital technology has gradually integrated into the real economy over the past two decades. As a major player in digital technology (Korinek and Stiglitz, 2021), China has witnessed a significant expansion in its digital economy. According to the statistics from the Chinese government, from 2008 to 2022, China's digital economy has witnessed remarkable growth, expanding from 4.8 trillion yuan to 50.2 trillion yuan, with its share in GDP rising from 15.2% to 41.5%. During this period, the revenue of Chinese SOEs has continued to increase, with the proportion of SOEs revenue to GDP reaching 68.3% in 2022. However, SOEs' market dynamics and motivation to engage in digitalisation are weak (Wu et al., 2021). The digitalisation level of SOEs is significantly lower than that of non-state-owned enterprises (as shown in Appendix Fig. A1). Many studies have confirmed that digitalisation enables high-quality development of enterprises (Acemoglu and Restrepo, 2018; Ghasemaghaei and Calic, 2019; Barrett et al., 2015). Given the importance of SOEs in developing countries such as China and the current state of digitalisation among SOEs, studying the motivations behind the digitalisation of SOEs is both representative and of practical significance.

The mixed ownership reform is a crucial aspect of the reform for Chinese SOEs. Previous studies have shown that the mixed ownership reform has significantly impacted firm financial performance (Gupta, 2005; Opoku-Mensah et al., 2022). However, the impact of mixed ownership reform on enterprise digitalisation has not been adequately explored. We construct a theoretical framework for the influence of the mixed ownership reform on digitalisation from the perspectives of its impact on compensation incentive mechanisms and risk attitude. Firstly, the mixed ownership reform improves the pay-performance sensitivities of management (Gupta, 2005). Digitalisation is key to enhancing firm performance and achieving high-quality development (Loebbecke and Picot, 2015). Management is incentivized to promote digitalisation and expects higher future compensations. From the risk perspective, the mixed ownership reform enables higher corporate risk-taking. digitalisation, as a risky investment practice that requires substantial resources (Scholz et al., 2020; Svahn et al., 2017), becomes more compelling when companies have higher risk-taking capabilities and tendencies. Therefore, the mixed ownership reform of SOEs may have a promoting effect on digitalisation.

Drawing upon the aforementioned theoretical framework, we utilized the word2vec word embedding tool to construct the measure of digitalisation and examine the impact of mixed ownership reform on it. The study finds that the mixed ownership reform of SOEs significantly promotes their digitalisation. Mechanism tests show that mixed ownership reform promotes digitalisation mainly by influencing the pay-performance sensitivities of management and corporate risk-taking. Heterogeneity analysis reveals that the effect is more significant in competitive industries and among enterprise decision-makers who did not experience the Great Chinese Famine during childhood.

In addition to the practical significance, the findings of our study offer significant contributions to the literature on digitalisation. Specifically, this study provides new perspectives and tools for measuring digitalisation. Existing literature has adopted various approaches to measuring it, such as surveys (Commander et al., 2010), investments in digital assets (Müller et al., 2018), and traditional text analysis (Chen and Srinivasan, 2024). However, these methods have limitations, including low accuracy, ostentatious investment, and subjective bias (Cycyota and Harrison, 2006; Triplett, 1999; Li et al., 2021a). Our study employs the Word2vec word embedding model to extract digitalisation-related information from unstructured data to alleviate these

issues. The Word2vec model, a deep learning algorithm in the machine learning domain, converts words into vectors based on contextual semantics, enabling synonym searching with vector algorithms. This approach not only overcomes the subjectivity inherent in manually constructed indicators in traditional text analysis but also ensures that the extracted key information on digitalisation originates from the research subjects, thereby enhancing the accuracy and validity of the indicator.

Besides, our study offers a novel perspective on the motivation of digitalisation. The drivers behind firms' digitalisation are multifaceted. First, economic motivations, such as the pursuit of monetary profit and improvements in financial performance (Verhoef et al., 2019; Müller et al., 2018), play a critical role. Second, managerial characteristics are also influential, including collaborative relationships, power allocation, and social capital (Matt et al., 2015; Marabelli and Galliers, 2017; Li et al., 2018). Third, policy and institutional factors, such as government investments in digital technology, tax incentives, and innovation subsidies, effectively promote digitalisation (Hess et al., 2016; Bharadwaj et al., 2013). While previous studies have predominantly focused on direct digitalisation support policies, few have examined the effects of indirect institutional arrangements, such as mixed ownership reform, on digitalisation. Our research concerns this gap.

Furthermore, our study contributes to the literature on mixed ownership reform. Prior research on mixed ownership reform has concentrated mainly on equity structure adjustments, managerial governance, and their effects on firm performance (Megginson et al., 1994; Megginson and Netter, 2001; Gupta, 2005; Opoku-Mensah et al., 2022), with limited attention to technological transformation. Our study extends the scope of mixed ownership reform research by examining its role in driving firms' digitalisation. Specifically, we uncover mechanisms through which such reform influences digitalisation processes by adjusting pay-performance sensitivities and enhancing firms' risk-taking capacity. Mixed ownership reform introduces more excellent market-oriented elements, significantly strengthening the correlation between managerial compensation and firm performance (Cai et al., 2018). Enhanced compensation incentives encourage managers to adopt long-term strategies, including digitalisation. This finding aligns with existing literature on the relationship between managerial incentives and corporate innovation (Lin et al., 2011). Moreover, mixed ownership reform optimizes equity structures and increases risk-taking, making firms more inclined to adopt emerging technologies and digitalisation. Our study validates these mechanisms, offering new insights into how mixed ownership reform influences technological transformation in firms.

The subsequent sections of the paper are structured as follows: Section "Literature review and research hypothesis" concisely reviews the relevant literature and presents the research hypotheses. Section "Data, variables, and descriptive statistics" details the data sources, variable construction process, and descriptive statistics. Section "Empirical Results and Analysis" presents the empirical results and analyses, including the empirical models, baseline regressions, robustness, endogeneity, and mechanism tests. Finally, we offer concluding remarks and implications of our findings.

## Literature review and research hypothesis

**The background of mixed ownership reform.** For a long time, the deficiencies of "lack of ownership" and "insider control" in Chinese SOEs have affected their operational efficiency and led to unfair distribution and rent-seeking corruption. These problems are expected to be alleviated with the mixed ownership reform by

introducing non-state property rights (Boardman and Vining, 1989; La Porta et al., 1999). The entry of private capital or foreign equity participation into SOEs can promote the flexibility of SOEs to switch their operating mechanisms and enhance the level of incentive compatibility of rights, responsibilities, and benefits internally, thus bringing about Pareto efficient outcomes. Furthermore, the mixed ownership shareholding structure introduces external shareholder supervision, which is helpful to make up for the previous lack of supervision and can effectively reduce the monitoring cost of the principal-agent relationships and curb the moral hazard behaviour of SOEs executives.

As a pillar of the Chinese economy, the government attaches great importance to reforming mixed ownership in SOEs. The work reports of the Chinese central government in 2021 and 2022 point to the need to deepen the reform of mixed ownership of SOEs. The 14th Five-Year Plan proposes that the mixed ownership reform will remain the direction of SOEs reform in the future (2021–2025). According to the public data, by the end of 2021, Chinese SOEs had introduced more than 2.5 trillion yuan of various types of private capital. The number of mixed ownership enterprises in central and local SOEs exceeds 70% and 54%, respectively, and most SOEs have carried out the mixed ownership reform.

**The impact of mixed ownership reform on digitalisation of the enterprise.** To begin with, concerning management incentives, the inefficiency observed in SOEs can be attributed to the absence of efficient monitoring and incentive mechanisms for managers, as discussed by Laffont and Tirole (1993). The causal relationship between the operational performance of SOEs and managerial behaviour is ambiguous, and factors such as government regulation constrain the incentive provisions in compensation contracts, leading to severe agency problems within the organization (Bardhan, 1997). As such, managers are not adequately motivated to strive for the development of the enterprise. On the other hand, non-state shareholders generally prioritize profit maximization as their business objective, thus having a strong motivation to enhance executive compensation. Firth et al. (2006) suggest that the relative change in managerial wealth compared to shareholder wealth can be used to measure the pay-performance sensitivities of the management team. Due to compensation constraints and alternative incentives such as political promotion, the pay-performance sensitivities in SOEs are relatively lower than in private enterprises (Kato and Long, 2006). Increasing the pay-performance sensitivities can promote the alignment of interests between principals and agents, alleviating the agency problem (Jensen and Murphy, 1990). This means that agents can achieve personal benefits while improving the operational performance of the enterprise.

The digitalisation of enterprise refers to the significant improvement of business models through integrating and utilizing digital technologies (Vial, 2019; Warner and Wäger, 2019). It leads to enhanced operational performance by fostering innovation, reducing transaction costs, improving internal governance, and increasing total factor productivity (Ghasemaghaei and Calic, 2019; Dana and Orlov, 2014; Rosenblat and Stark, 2016; Commander et al., 2010; Loebbecke and Picot, 2015). Acemoglu and Restrepo (2018) found that firms' adoption of digital technologies promotes automation and intelligence in production operations, leading to increased productivity. Digitalisation also facilitates information dissemination, thereby reducing transaction search costs and contract costs (Dana and Orlov, 2014). Further mixed ownership reform and digitalisation enable optimized corporate governance through data mining and data-driven intelligent decision-making, leading to more sophisticated and scientific enterprise management. Collecting customer

interaction data in products and services creates new business value (Porter and Heppelmann, 2014). Li and Yu (2015) discovered that the involvement of non-state shareholders in governance could lead to enhancements in the compensation incentive policies for senior executives within SOEs, motivating managers to work diligently and promoting increased investment in technology activities. We believe that the mixed ownership reform can effectively enhance the pay-performance sensitivities of the management team. Since digitalisation can improve operational performance and profitability, the increased pay-performance sensitivities induced by the mixed ownership reform will further motivate managers to promote digitalisation. Based on our analysis, we put forward the following hypothesis:

H1: The mixed ownership reform of SOEs promotes digitalisation by increasing the pay-performance sensitivities of the management team.

Second, in terms of risk-taking, as digitalisation requires companies to reinvent their business models, the process will expose companies to uncertainty shocks and disruptive risks (Scholz et al., 2020). Path dependencies such as corporate inertia, available resources, and capabilities are difficulties that companies need to overcome for digitalisation (Svahn et al., 2017). For example, having highly optimized but rigid traditional production processes makes resources not easily reallocated (Kohli and Johnson, 2011), and managers and employees may show antagonistic attitudes when disruptive technologies are introduced into the enterprise (Singh and Hess, 2017). If firms are less willing to take risks, according to prior research by John et al. (2008), they prefer conservative investment strategies. In other words, firms may be reluctant to promote digitalisation when they are less willing to take risks. The SOEs typically carry a heavy policy burden to undertake many social functions for the government (Lin et al., 1998). They usually have low risk-taking levels because excessive risk-taking is not conducive to achieving the government's political goals and social stability functions. Therefore, firms subject to government intervention prefer conservative investment strategies (Fogel et al., 2008). Shleifer and Vishny (1994) show that the reform of firm ownership from state-owned to private makes it more costly for politicians to intervene in firms. The political pressure exerted by the government on firms to achieve political goals is reduced. This may cause firms to shift their conservative business strategies. At the same time, privatization alleviates the agency problem of SOEs and gives firms a greater incentive to improve their corporate governance mechanisms, raising their risk-taking level (Boubakri et al., 2013). Therefore, the increased risk-taking level of SOEs following the mixed ownership reform will improve the undertaking of digitalisation activities. In summary, we posit the following hypothesis:

H2: The mixed ownership reform of SOEs promotes digitalisation by increasing corporate risk-taking.

We conduct empirical tests using data from Chinese A-share listed SOEs to validate the theoretical analysis and research hypotheses. To enhance the accuracy of the research findings, we employ a word vector model to refine the digitalisation measure and characterize mixed ownership reform in SOEs from two dimensions: equity structure and top governance structure. To ensure the robustness of the conclusions, we supplement the baseline regression with a series of robustness tests, endogeneity tests, and heterogeneity analyses. Furthermore, we address the proposed hypotheses through mechanism tests. The following sections provide a detailed discussion of the data sources, measurement construction, and empirical results.

## Data, variables, and descriptive statistics

**Data.** For our research sample, we choose Chinese A-share listed SOEs spanning from 2010 to 2020 and proceed with the data

processing as follows: (1) We exclude financial and insurance, ST and ST\* listed companies<sup>1</sup>; (2) exclude the units with missing observations; The final sample has 10445 observations. The financial data used in this study are sourced from the China Stock Market & Accounting Research Database (CSMAR), and the data of related companies' annual reports are obtained from the official websites of the Shenzhen Stock Exchange and Shanghai Stock Exchange.

### Variable construction

*The mixed ownership reform.* We assess the level of mixed ownership reform through two distinct dimensions: equity structure and top governance structure. To quantify the equity structure, we utilize the ratio of non-state shareholder equity among the top ten shareholders, denoted as *NonSOEs\_Equ*. Additionally, to assess the top governance structure, we calculate the proportion of directors appointed by non-state shareholders to the total number of directors on the company's board, denoted as *NonSOEs\_Dir*. The shareholder-appointed directors are determined here following Cheng and Wei (2013): In instances where a shareholder is a natural person, we consider them as having been appointed as a director if they currently hold a directorship position within the listed company. Regarding non-state-owned enterprise shareholders, the judgment standard is whether the directors of the listed company hold positions in the non-state-owned shareholders<sup>2</sup>. We can obtain more robust and prosperous conclusions by measuring the degree of mixed ownership reform from the above two dimensions.

*The digitalisation of enterprise.* Loughran and McDonald (2011) showed that quantifying and analyzing key information through word frequency can reflect business characteristics. With the advancement of natural language processing technology, using text analysis to construct the enterprise digitalisation metric is increasingly becoming a major approach (Chen and Srinivasan, 2024; Chen et al., 2024; Wu and Li, 2024; Wu et al., 2021; Yuan et al. 2021; Zhao et al. 2021). For example, Chen and Srinivasan (2024) use textual descriptions of digitalisation in business discussion texts in annual reports (10-K reports) as measurement; Wu et al. (2021) refer to the digital transformation-related literature, essential policy documents, and research reports to compile a library of keywords for digital transformation, and measure enterprises' digitalisation based on the number of occurrences of the words in texts of listed companies. However, there are four main problems in the current study. First, there is no consistent definition of keyword selection in the digital transformation lexicon, and the study by Wu et al. (2021) is a pioneering and influential study that uses the lexicon method to measure digital transformation in the context of Chinese annual reports. The study summarizes the keywords of digital transformation from the perspectives of "underlying technology use" and "practical application of technology", but the above categorization concepts are vague, with low precision and insufficient completeness. For example, "Business Intelligence" is included in the application of "underlying technology", while "Intelligent Marketing" is included in the "practical application of technology". The Internet of Things is included in cloud computing. Considering digital transformation only as an application based on cloud computing, big data, artificial intelligence, and blockchain technologies, Vial (2019) and Bharadwaj et al. (2013) reveal that the application of a wide range of technologies such as information technology, computing, communication, and connectivity is the foundation of digital transformation. In addition, Yuan et al. (2021) used the descriptions of digital transformation in China's policy documents at a specific time to construct an

enterprise key dictionary. Zhao et al. (2021) took some samples of enterprises' annual reports to find the key information and build the dictionary in specific areas, namely, the application of digital technology, Internet business models, intelligent manufacturing, and modern information systems. The keyword dictionaries and indicators obtained in various ways in the past are different. The keyword sources, dictionary construction ideas, and choices are relatively subjective and lack theoretical support. Given that the application of digital technology is the foundation of enterprise digital transformation (Verhoef et al., 2019; Singh and Hess, 2017), we start with the definition of digital transformation and extract the key textual information of digital transformation from three perspectives. Second, the completeness of keywords is insufficient. It is difficult to exhaust digital transformation keywords manually. In the case of a long textual period, the variety of ways textual content is expressed in different temporal contexts makes it more challenging to exhaust keywords. Third is the problem of subjective bias. Previous research methods relied entirely on manual techniques to select keywords, and such a process cannot avoid bias caused by individual capabilities and subjective factors. Fourth, there is a lack of keyword validity. The keywords selected by the study may not appear or appear too infrequently in the context of annual reports. Some studies select keywords through external literature, working reports, and policy texts. However, Jiang et al. (2021) found that the expression of words in annual reports has more complex characteristics compared to the written expression of words in literature and policy texts. That is, the enterprise digital transformation lexicon constructed based on literature and policy texts may not effectively reflect the expression characteristics of digital transformation in the context of the annual report. In a subsequent study, we show many expressions related to digitalisation words based on specialization, English, and even acronyms based on the context of the enterprise's annual report<sup>3</sup>. In contrast, relatively few policy texts allow the above expressions. To summarize, the dictionary method based on traditional text analysis faces problems such as lack of completeness, subjective bias, significant keyword variation, and lack of validity. These problems affect the accuracy of the results of digital transformation measurement and quantitative analysis.

We refer to previous studies (Li et al., 2021a; 2021b) to improve the traditional manual keyword method with the help of the unsupervised machine learning model of Word2vec, a landmark technology in the field of similar word expansion (LeCun et al., 2015). The Word2vec model is based on the well-recognized concept in linguistics that words occurring in similar contexts have similar meanings. It converts words into vectors based on their contextual contexts so that word similarity can be computed based on the vector similarity and thus learns to predict all of its approximate words. In addition, the Word2vec technique requires specific text to train the model, and according to the research needs, we choose the content of managerial discussion and analysis (MD&A) disclosed in the annual reports of listed companies as the research text. This text reflects the review of business conditions and the outlook for future development of the enterprise during the reporting period, and it has been widely recognized as the research object (Loughran and McDonald, 2011).

Combining previous studies and definitions of enterprise digitalisation, we look for expressions of enterprise digitalisation in three dimensions: first, broad and traditional digitalisation seed words, referring to studies such as Vial (2019), we use "Internet", "Information technology" and "Communication technology" as seed words; second, cutting-edge digitalisation seed words. In the definition of digitalisation, one strand of literature highlights that enterprise digitalisation is based on cutting-edge digital



technologies. For example, Singh and Hess (2017) argue that new digital technologies refer to social media, mobile connectivity, analytics or embedded devices, etc.; Warner and Wäger (2019) argue that new technologies refer to artificial intelligence, cloud, blockchain, and IoT. Combined with the context of Chinese expression, we put “Cloud computing”, “Big data”, “Artificial intelligence”, “Blockchain”, “Internet of Things”, and “Mobile Internet” as the seed words; thirdly, we refer to the use of digital technology seed words in a general way, and we use the words “Digital transformation”, “Digitalisation” and “Digital technology” as seed words. Finally, the above three parts of words are aggregated to get the enterprise digitalisation seed words. Given the characteristics of digitalisation definition, we also use the seed words based on the above frontier digital technologies to construct a dictionary and indicator of digitalisation based on frontier technologies, which are used for robustness testing.

After obtaining the seed words, we use the Word2vec model to output similar words in the MD&A text of listed companies. Following the methodology of Li et al. (2021a, 2021b), we employ the CBOW (Continuous Bag of Words) model within Word2vec for training, setting the vector dimension to 300 and outputting the top 100 similar words for each seed word based on similarity scores. This process involves two key steps: converting words into vectors and estimating the similarity between the word vectors. We use the CBOW model to estimate the word vectors. The core idea of the CBOW model is to predict the probability of the current word based on its context, and by maximizing the likelihood function, the corresponding vectors for each word are obtained. Mathematically, this can be expressed as follows:

$$\text{Vector\_word} = \max_{w \in T} \sum \log p(w|\text{Text}(w)) \quad (1)$$

Note:  $T$  represents the corpus,  $w$  represents a word, and  $\text{Text}(w)$  denotes the context.

Second, the similarity between word vectors is measured using cosine similarity. Based on the word vectors derived from the CBOW model, we calculate the cosine similarity between vectors to determine their similarity. In other words, cosine similarity allows us to identify words similar to the seed words. For two given word vectors  $X$  and  $Y$ , the cosine similarity is computed as follows:

$$\text{Similarity} = \cos(\theta) = \frac{X \times Y}{|X| \times |Y|} = \frac{\sum_{i=1}^n X_i \times Y_i}{\sqrt{\sum_{i=1}^n X_i^2} \times \sqrt{\sum_{i=1}^n Y_i^2}} \quad (2)$$

We output each seed word's top 100 similar words based on their cosine similarity scores. For instance, using the seed word “Digital Technology”, Table 1 presents the top 10 most similar words in the Chinese context and their similarity scores.

The results in Table 1 indicate that most of the words output by the word vector model are highly relevant to “Digital technology”, demonstrating the validity of the research method adopted and highlighting the efficiency and accuracy of using machine learning to identify similar words compared to traditional manual methods. However, there may still be a small number of irrelevant words among the outputs. Following Li et al. (2021a, 2021b), we manually verify the machine learning results and exclude a small number of words unrelated to digitalisation (e.g., “Modern Technology”, “Cultural Media”, “Multimedia Technology”, as shown in Table 1). Ultimately, through this process, we developed a digitalisation dictionary based on the MD&A text from annual reports spanning 2010 to 2020.

We extracted the aggregate frequency of keywords in the MD&A text of listed companies' annual reports through the dictionary. Concerning previous studies (Wu et al., 2021), we logarithmized it to get the enterprise digitalisation indicator

**Table 1 Example of similar words for “Digital Technology” output by Word2vec.**

No.	Seed word	Similar words	Similarity
1	Digital Technology	Information Technology	0.504
2	Digital Technology	Software Technology	0.498
3	Digital Technology	Communication Technology	0.466
4	Digital Technology	Modern Technology	0.408
5	Digital Technology	Cultural Media	0.406
6	Digital Technology	Computer Technology	0.402
7	Digital Technology	Multimedia Technology	0.391
8	Digital Technology	Artificial Intelligence	0.388
9	Digital Technology	Mobile Internet	0.380
10	Digital Technology	E-commerce	0.377

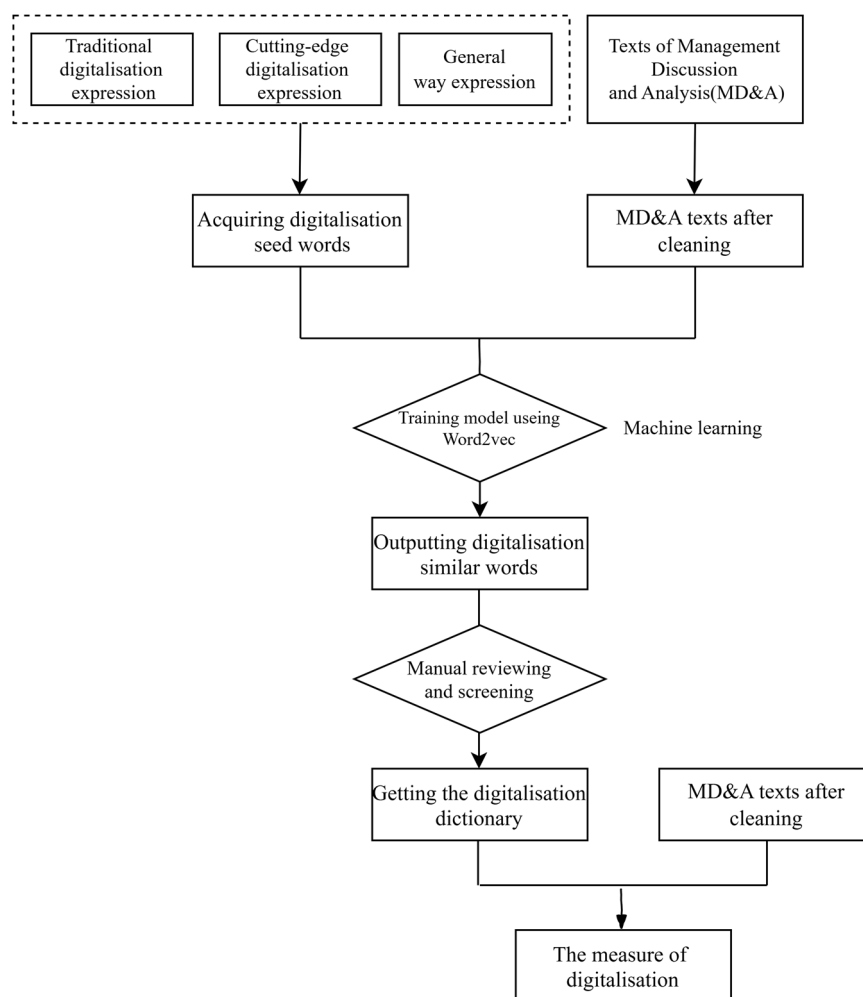
(Digital). Figure 1 shows the process of building the digitalisation measure using machine learning.

### Validation of digitalisation of enterprise measure

**Geographical Characteristics.** We calculate the digitalisation mean value of all listed companies in any city during the sample period on a city-by-city basis and obtain the digitalisation geographic distribution map of listed companies in China in Fig. 2. The map presents the following two distinctive features. First, the digitalisation level of enterprises in the eastern region, especially in the coastal region, is significantly higher than that in the central, western, and northeastern regions. Second, within the same provincial area, the digitalisation level of provincial capital cities is generally higher than in other regions. The main reason is that the digital infrastructure and economic development of the eastern region and provincial capital cities are higher than those of their western, northeastern, and non-provincial capital regions, which is conducive to the digitalisation of enterprises in the regions mentioned above. In this paper, the constructed indicators reflect the actual situation regarding the geographical characteristics of enterprise digitalisation.

**The time trend.** The two curves in Fig. 3 represent our constructed graph of the time trend of the annual mean value of digitalisation of listed companies and the size of China's digital economy as a share of GDP. To make the graph more intuitive, we multiply the proportion of China's digital economy size to GDP by 10. The graph reflects two features: first, enterprise digitalisation as a whole shows a growing trend, and from 2010 to 2020, the proportion of China's digital economy size to GDP grew from 18.6% to 38.6%, and the mean value of enterprise digitalisation during the same period increases from 0.942 roses to 2.713, and the above indicators all show a year-on-year increasing trend, indicating that the growth of enterprise digitalisation is consistent with the overall trend of China's digital economy development; secondly, enterprise digitalisation showed a more apparent upward leap in 2015, a finding that is in line with the viewpoints of Peng et al. (2023). The main reason is that the Chinese government stepped up its efforts to promote the digitalisation of the real economy in that year. 2015 was the first year that Chinese leaders proposed the construction of “Digital China”. Also, the State Council of China proposed the implementation of the National Big Data Strategy and issued documents such as the “Action Plan for the Promotion of Big Data Development” and “Made in China 2025” to promote the rapid development of digitalisation. The measure reflects the actual situation from the time trend of digitalisation.

**The definition of variables.** Based on previous studies, we specify several variables of firm and management characteristics as



**Fig. 1 The process of building the digitalisation measure.** This figure illustrates the steps involved in constructing a measure for digitalisation. The arrow symbols represent processes.

control variables. Table 2 presents the names and corresponding definitions of each variable used in the study.

**Descriptive statistics.** Table 3 displays the descriptive statistics. The mean value of digitalisation level (*Digital*) of SOEs is 1.659; the variance is 1.328, the maximum value is 6.275, and the minimum value 0. This indicates significant differences in the digitalisation levels of different enterprises. For the indicators of SOEs' mixed ownership reform, the average proportion of directors appointed by non-state shareholders is only 3.8%, which indicates the overall low level of non-state shareholders' participation in governance. The proportion of non-state shareholder equity among the top ten shareholders is 13.9%, reflecting the overall low shareholding ratio of non-state shareholders at the level of key shareholders.

### Empirical Results and Analysis

The empirical model is as follows:

$$Digital_{i,t} = \beta_0 + \beta_1 NonSOEs\_Dir_{i,t} / NonSOEs\_Equ_{i,t} + \beta_2 Controls_{i,t} + \sum Industry + \sum Year + \varepsilon_{i,t} \quad (3)$$

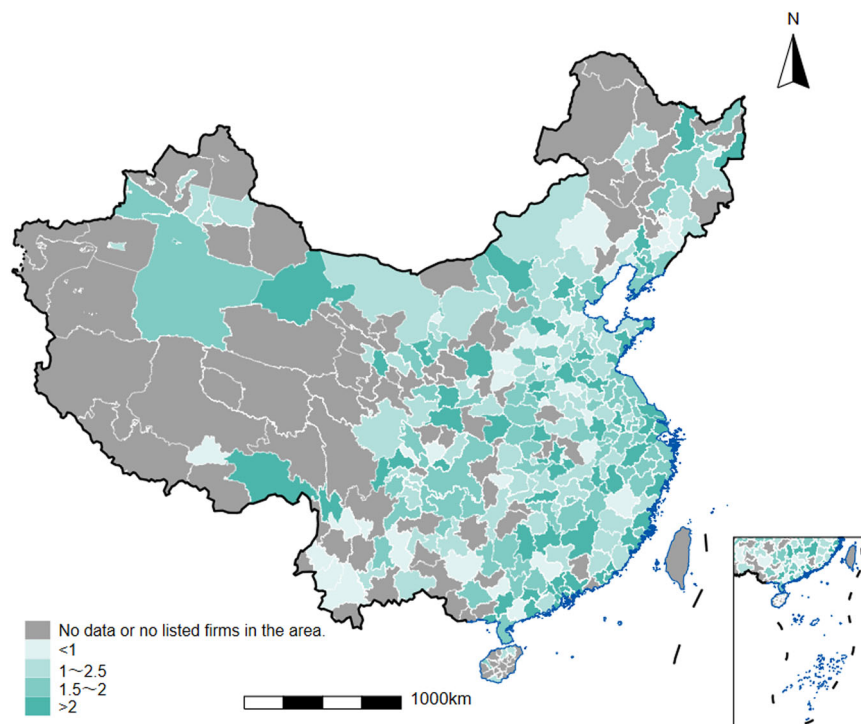
$Digital_{i,t}$  is the dependent variable that captures the digitalisation of firm  $i$  in year  $t$ . The independent variable  $NonSOEs\_Dir_{i,t}$  and  $NonSOEs\_Equ_{i,t}$  assessed the mixed ownership reform from the perspectives of top governance structure and equity structure, respectively.  $Controls_{i,t}$  represents control variables. We also control for industry and year fixed effects.

**Baseline regression.** Table 4 shows the baseline regression results. Columns 1 and 3 do not include control variables and the only control for time and industry fixed effects and show that the impact of the proportion of non-state shareholder directors and the proportion of non-state shareholder ownership among the top ten shareholders on the digitalisation of enterprise is significantly positive at the 1% level. Columns 2 and 4 show that these two mixed ownership reform variables remain significant after including control variables. This implies that the mixed ownership reform of SOEs will significantly promote the digitalisation of enterprises.

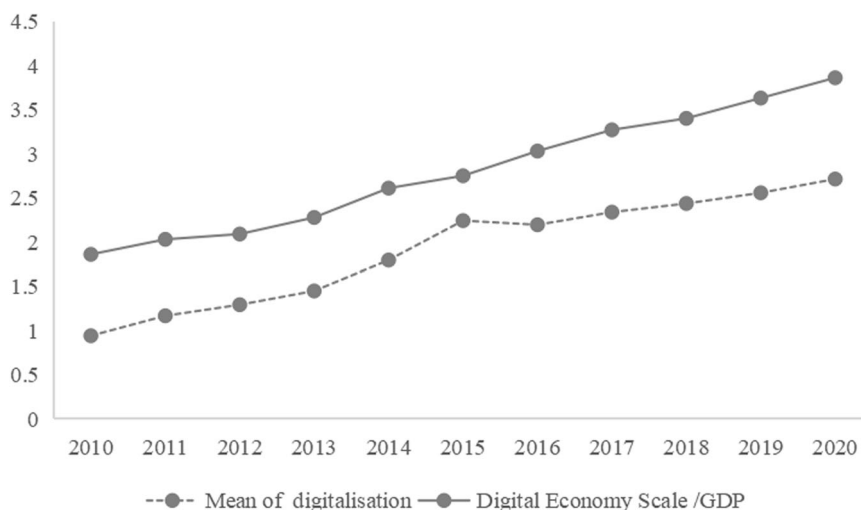
### Robustness tests

**Estimating with alternative independent variables and regression models.** First, we use the logarithm of the number of directors appointed by non-state shareholders in the board of directors ( $LnNonSOEs\_Dir$ ). The regression results reported in column 1 of Table 5 shows that the alternative mixed reform indicators continue to contribute positively and significantly to digitalisation. In addition, due to the presence of zero values in the dependent variable *Digital*, we further employ the Tobit model for the regression, and the results reported in columns 2 and 3 show that our basic conclusions remain robust.

**Re-specifying the dependent variable.** We attempt to reconstruct the dictionary of enterprise digitalisation by considering the definition of digital technology. Some studies suggest that the digital technology



**Fig. 2 The geographical characteristics of digitalisation in China.** This figure calculates the mean digitalisation level of all Chinese listed firms in each city during the sample period. Darker colors indicate a higher level of digitalisation among listed firms within the city. It reflects the higher level of digitization in economically developed regions.



**Fig. 3 The time trend of digitalization.** In this figure, the horizontal axis represents the year, and the vertical axis shows the annual mean of digitalisation and the value of the proportion of China's digital economy scale to GDP multiplied by 10. The gray solid line (Digital Economy Scale/GDP) reflects the scale of the digital economy, while the gray dashed line (mean of digitalisation) represents the level of corporate digitalisation. This figure indicates a significant increase in digitalisation in 2015, which can be attributed to a series of concentrated and large-scale policies introduced by the Chinese government to support corporate digitalisation during that year, such as the construction of "Digital China", the implementation of the national big data strategy, the publication of the "Action Plan for the Promotion of Big Data Development" and "Made in China 2025".

employed for digitalisation is an advanced form of technology, distinct from traditional information and communication technologies (Warner and Wäger, 2019). Given this, we reconstruct the digitalisation dictionary and indicator (Digital\_RE) as a robustness test. In this dictionary, keywords such as "internet technology", "information technology", and "web technology", which refer to traditional digital technologies, are discarded. The findings in columns 1 and 2 of Table 6 demonstrate the robustness of the mixed ownership reform in fostering the digitalisation of enterprises.

*Propensity score matching (PSM).* we also utilize the PSM method for robustness tests. The entire sample is grouped according to the presence or absence of non-state shareholder-appointed directors (*NonSOEs\_Dummy*). The variables of *Size*, *Age*, *Leverage*, *Growth*, *TBQ*, *ROA*, *ROA*, *Indep*, *Board*, and *Dual* are used as covariates to identify a control group with similar characteristics to the treatment group, using the nearest neighbour matching, radius matching, and kernel matching methods, respectively. The balance tests were performed on all covariates before PSM

**Table 2 The definition of variables.**

Variable	Definition
Digital	The degree of digitalisation of enterprise is measured by summing the frequency of digital keywords in MD&A text plus 1 and taking the natural logarithm.
NonSOEs_Dir	The ratio of the number of directors appointed by non-state shareholders to the total number of directors in the company.
NonSOEs_Equ	The percentage of all non-state shareholders' equity among the top ten shareholders.
Size	The natural logarithm of the size of the company's assets.
Age	The number of years since the company was founded.
Leverage	The dividing of the firm's total liabilities by its total assets.
Growth	The company's income growth rate in year t over year t-1.
ROA	The net income after taxes is divided by total assets.
Indep	The ratio of the number of independent directors in the company to the total number of directors on the board.
Board	The number of board members taking the natural logarithm.
Dual	The dummy variable for whether the chairman is also the CEO is 1 if he is, and 0 otherwise.

**Table 3 Descriptive statistics.**

Variable	Mean	SE	Median	Min	Max
Digital	1.659	1.328	1.609	0.000	6.275
NonSOEs_Dir	3.796	9.118	0.000	0.000	85.714
NonSOEs_Equ	0.139	0.149	0.082	0.000	0.951
Size	8.847	1.412	8.680	6.203	12.912
Age	18.130	6.047	18.000	0.000	54.000
Leverage	0.506	0.202	0.514	0.078	0.937
Growth	0.087	0.241	0.037	-0.489	1.083
ROA	0.052	0.058	0.047	-0.156	0.237
Indep	0.370	0.058	0.333	0.000	0.800
Board	2.207	0.198	2.197	0.693	2.890
Dual	0.093	0.290	0.000	0.000	1.000

**Table 4 Baseline regression results.**

VARIABLES	(1) Digital	(2) Digital	(3) Digital	(4) Digital
NonSOEs_Dir	0.010*** (3.69)	0.011*** (3.94)		
NonSOEs_Equ			0.368*** (2.82)	0.243* (1.85)
Constant	1.621*** (59.42)	0.607 (1.50)	1.609*** (50.06)	0.709* (1.75)
Observations	10,445	10,445	10,441	10,440
R-squared	0.343	0.360	0.340	0.355
Controls	YES	YES	YES	YES
IND FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Note: The adjusted for clustering robust standard errors are in parentheses. Standard errors are clustered at the firm level, and significance levels are denoted as \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .

regression estimation was executed. The differences in characteristics between firms with or without directors appointed by non-state shareholders were effectively eliminated after PSM. The post-matching estimation results presented in Table 7 remain consistent with the baseline regression results, demonstrating that the study findings remain robust after addressing the problem of sample selection bias.

**Endogeneity test.** Referring to Cai et al. (2018) and Fan et al. (2013), we utilize the historical concessions in China as the instrumental variable of mixed ownership reform, selected for its relevance, exogeneity, and exclusivity. And we give the following reasons:

**Table 5 Changing the mixed ownership reform measurement and using the Tobit model.**

VARIABLES	(1) Digital	(2) Digital	(3) Digital
LnNonSOEs_Dir	0.261*** (4.48)		
NonSOEs_Dir		0.013*** (3.97)	
NonSOEs_Equ			0.261* (1.70)
Constant	0.652 (1.62)	-1.203** (-2.44)	-1.108** (-2.24)
Observations	10,445	10,445	10,440
R-squared	0.361	0.118	0.116
Controls	YES	YES	YES
IND FE	YES	YES	YES
Year FE	YES	YES	YES

Note: The adjusted for clustering robust standard errors are in parentheses. Standard errors are clustered at the firm level, and significance levels are denoted as \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .

**Table 6 Changing digitalisation measurement.**

VARIABLES	(1) Digital_RE	(2) Digital_RE
NonSOEs_Dir	0.010*** (3.76)	
NonSOEs_Equ		0.190* (1.67)
Constant	0.085 (0.24)	0.169 (0.48)
Observations	10,445	10,440
R-squared	0.350	0.345
Controls	YES	YES
IND FE	YES	YES
Year FE	YES	YES

Note: The adjusted for clustering robust standard errors are in parentheses. Standard errors are clustered at the firm level, and significance levels are denoted as \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .

Before the First Opium War, China's Qing Dynasty adopted a policy of isolation, severing interactions with other countries. Following the war, a series of agreements and treaties required the Chinese government to open specific trading ports and establish concessions, allowing foreign capital to invest, establish factories, conduct missionary work, and promote economic and cultural exchange. Similar to colonisation, establishing concessions may



**Table 7 The PSM test.**

VARIABLES	(1) Digital_Neighbor	(2) Digital_Radius	(3) Digital_Kernel
NonSOEs_Dummy	0.007** (2.47)	0.011*** (3.84)	0.011*** (3.86)
Constant	1.236** (2.19)	0.705* (1.80)	0.701* (1.80)
Observations	3365	10,063	10,159
R-squared	0.380	0.361	0.362
Controls	YES	YES	YES
IND FE	YES	YES	YES
Year FE	YES	YES	YES

Note: The adjusted for clustering robust standard errors are in parentheses. Standard errors are clustered at the firm level, and significance levels are denoted as \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .

**Table 8 Instrumental variable tests.**

VARIABLES	(1) Phase I NonSOEs_Dir	(2) Phase II Digital	(3) Phase I NonSOEs_Equ	(4) Phase II Digital
NonSOEs_Dir		0.073** (2.09)		
NonSOEs_Equ				5.499* (1.93)
Concession	1.747*** (3.64)		0.023*** (3.21)	
Kleibergen-Paap rk LM statistic	13.216***			10.322***
Kleibergen-Paap rk Wald F statistic	13.243			10.317
Observations	10,445	10,445	10,440	10,440
R-squared	0.062	0.191	0.060	0.032
Controls	YES	YES	YES	YES
IND FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Note: The adjusted for clustering robust standard errors are in parentheses. Standard errors are clustered at the firm level, and significance levels are denoted as \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .

have long-term effects on developing local markets and legal institutions (Acemoglu et al., 2001). Concession areas were characterised by an emphasis on private enterprise, property rights, and market economies, resulting in a more developed institutional foundation for SOEs reforms in these regions, thus satisfying the relevance requirement for instrumental variables.

However, the existence of concessions is a historical legacy, making this variable highly exogenous. Its influence is mainly reflected in shaping local market mechanisms (Fan et al., 2013). Digitalisation, on the other hand, is largely a product of modern information technology development. This disconnect makes it unlikely that the history of concessions directly affects contemporary corporate digitalisation. Digitalisation relies more on modern information infrastructure, corporate economic incentives, and investment capacity (Vial, 2019; Bharadwaj et al., 2013), which are not directly related to the historical presence of concessions in the firm's location. Thus, while concessions do not directly drive modern corporate digitalisation, they may exert an indirect effect through their impact on mixed ownership reform. This satisfies the exclusion restriction requirement for instrumental variables. Finally, we construct an instrumental variable (*Concession*), which takes the value of 1 if the listed company is located in a province that has historically established a concession area and 0 otherwise.

Table 8 reports the instrumental variable regression results. Column 1 indicates that the instrumental variable has a significant positive effect on the proportion of directors appointed by non-state shareholders (*NonSOEs\_Dir*) and passed the weak instrumental variables test and rejection of the under-identification hypothesis. The two-stage regression results in column 2 indicate that the impact of the proportion of directors appointed by non-state shareholders on the digitalisation of enterprise remains positively significant. Similarly, column 3 indicates that the instrumental variable has a significant positive effect on the proportion of non-state shareholder equity among the top ten shareholders (*NonSOEs\_Equ*), and it passes the weak instrumental variable test as well as the under-identification hypothesis. The two-stage regression results in column 4 indicate that the effect of the proportion of non-state shareholder equity among the top ten shareholders on digitalisation of enterprise is still positively significant.

**Mechanism tests.** The literature shows that mixed ownership reform can enhance management pay-performance sensitivities and form a performance-oriented incentive mechanism. The digitalisation of enterprise is an essential means to promote corporate performance, and driving digitalisation will give management higher incentives and performance expectations in the future. In China, monetary compensation is the most critical component of managerial compensation incentives, and the assessment of the relationship between managerial compensation and performance in major companies focuses on the relationship between monetary compensation and performance (Cai et al., 2018). Referring to Firth et al. (2006) and Cai et al. (2018), we measured managerial compensation by taking the natural logarithm of the average compensation of the top three management personnel disclosed by the company. We used the ratio of management compensation to the natural logarithm of company net profit to measure the management pay-performance sensitivities (*Lnpay*). The regression results in columns 1 and 2 of Table 9 show that *NonSOEs\_Dir* and *NonSOEs\_Equ* significantly increase the pay-performance sensitivities of management, indicating that the pay-performance sensitivities are an essential mechanism for the mixed ownership reform to drive digitalisation. Thus, Hypothesis H1 is confirmed.

Second, the literature analysis also shows that digitalisation is a risky activity, and Chinese SOEs tend to have low risk-taking levels due to their political objectives and social stability functions. The mixed ownership reform can reduce the government's influence on corporate investment decisions, and firms will reorient toward the goal of maximizing corporate value and choosing more aggressive business strategies, undertaking a higher level of corporate risk-taking, which in turn promotes digitalisation. We refer to the research design of John et al. (2008) and use the level of firm earnings volatility (*Risk\_taking*) over a specific period to measure the extent of firm risk-taking. We adopt a 3-year observation period (from  $t$  years to  $t + 2$  years) due to the typical tenure of executives in Chinese listed firms, which is generally 3 years. The procedure of *Risk\_taking* calculation is shown in Eqs. (4) and (5), and the results shown in columns 3 and 4 of Table 9 indicate that the appointment of directors by non-SOE shareholders and the proportion of non-SOE shareholders equity among the top ten shareholders of the firm significantly enhance the firm's risk-taking. The results indicate that corporate risk-taking is an important mechanism through which SOEs' mixed ownership reform promotes the digitalisation of enterprises. Thus, Hypothesis H2 is confirmed.

$$\text{Risk\_taking}_{i,j} = \sqrt{\frac{1}{n-1} \sum_{j=1}^n \left( \text{Adj\_Roar}_{i,j} - \frac{1}{n} \sum_{j=1}^n \text{Adj\_Roar}_{i,j} \right)^2}, n = 3 \quad (4)$$

Table 9 Mechanism tests: the mixed ownership reform, pay-performance sensitivities, and corporate risk-taking.				
VARIABLES	(1)	(2)	(3)	(4)
	Lnpay	Lnpay	Risk_taking	Risk_taking
NonSOEs_Dir	0.001*** (2.92)		0.001** (2.02)	
NonSOEs_Equ		0.059*** (3.65)		0.140*** (3.63)
Constant	2.408*** (56.50)	2.415*** (57.35)	0.619*** (7.44)	0.631*** (7.62)
Observations	9432	9428	9434	9433
R-squared	0.640	0.640	0.146	0.149
Controls	YES	YES	YES	YES
IND FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Note: The adjusted for clustering robust standard errors are in parentheses. Standard errors are clustered at the firm level, and significance levels are denoted as \*\*\*p < 0.01, \*\*p < 0.05, \*p < 0.1.

Table 10 Heterogeneity test: degree of industry competition.				
VARIABLES	(1)	(2)	(3)	(4)
	Digital Compete=1	Digital Compete=0	Digital Compete=1	Digital Compete=0
NonSOEs_Dir	0.011*** (3.10)	0.007** (2.06)		
NonSOEs_Equ			0.345* (1.76)	-0.027 (-0.17)
Constant	0.347 (0.55)	0.061 (0.13)	0.437 (0.69)	0.106 (0.23)
Observations	6032	4413	6030	4410
R-squared	0.327	0.417	0.322	0.416
Controls	YES	YES	YES	YES
IND FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Empirical p-value	0.019		0.000	

Note: The adjusted for clustering robust standard errors are in parentheses. Standard errors are clustered at the firm level, and significance levels are denoted as \*\*\*p < 0.01, \*\*p < 0.05, \*p < 0.1.

$$Adj\_Ro a_{i,j} = \frac{EBIT_{i,j}}{ASSET_{i,j}} - \frac{1}{m_j} \sum_{k=1}^m \frac{EBIT_{k,i}}{ASSET_{k,j}}$$
 (5)

Note: *Roa* is the return on assets, *Adj\_Roa* is the return on assets adjusted by industry mean, *EBIT* is Earnings before interest and tax, *ASSET* is total company assets; *i* denotes the *i* company, *j* denotes the *j* year of the forward-looking 3 years; *m<sub>j</sub>* is the total number of companies in industry *m* in the forward-looking *j* year, and *k* is the *k* company in industry *m*.

**Heterogeneity analysis.** Jiang (2022) argues that the most crucial role of heterogeneity analysis lies in further testing the mechanisms. Based on this rationale, we examine the mechanisms of the pay-performance sensitivities of management and risk-taking through industry competition levels and decision-makers’ early experiences, respectively.

**Degree of industry competition.** Competitive SOEs face significant competitive pressure, necessitating the design of more reasonable compensation contracts to maximize shareholder interests and enhance overall corporate performance. Compared to monopolistic SOEs, mixed ownership reform is expected to have a

more pronounced effect on improving the pay-performance sensitivities of management in competitive SOEs. To test this hypothesis, we refer to Ke et al. (2017) and divide the sample into regulated and competitive industries. The variable *Compete* is assigned a value of 1 for firms in competitive industries and 0 otherwise, and a grouped regression analysis is performed. As shown in Columns (1) and (3) of Table 10, the coefficients for *NonSOEs\_Dir* and *NonSOEs\_Equ* are significantly positive in competitive industries. Conversely, in regulated industries, the significance and magnitude of the *NonSOEs\_Dir* coefficient decrease in Column (2), while the coefficient for *NonSOEs\_Equ* in Column (4) is insignificant. An inter-group correlation test confirms that the differences between these coefficients are statistically significant. These findings indicate that the positive effects of mixed ownership reform on digitalisation are more pronounced in competitive industries. This may be attributed to the stronger impact of mixed ownership reform on enhancing the pay-performance sensitivities of management in these industries.

**Corporate decision-makers who experienced famine during childhood.** Previous research has found that corporate decision-makers’ childhood experiences of famine can shape deeply rooted and unchangeable risk perceptions (Wang et al., 2022). If mixed ownership reform influences corporate digitalisation by affecting firms’ risk-taking, the effect should be more pronounced in firms that are more susceptible to such influences. In China, the chairman of the board significantly impacts the firm’s operational decisions (Feng and Johansson, 2018). We examine the risk-taking mechanism based on the historical event of the Great Chinese Famine and its impact on shaping the risk perceptions of the chairman. Referring to the study by Xu and Li (2016), the variable *Famine* is assigned a value of 1 if the chairman’s birth year falls between 1947 and 1961 (indicating that they experienced famine during childhood), and 0 otherwise, and a grouped regression analysis is conducted. The regression results in Columns (2) and (4) of Table 11 show that in the group where *Famine* = 0, the coefficients for *NonSOEs\_Dir* and *NonSOEs\_Equ* are significantly positive, whereas in the group where *Famine* = 1, the coefficients are not significant. An inter-group correlation test confirms that the differences between these coefficients are statistically significant. These findings suggest that mixed ownership reform has a more pronounced effect on promoting digitalisation in firms led by decision-makers without childhood famine experiences. This may be related to the fact that mixed ownership reform more effectively enhances risk-taking in such firms.

Conclusion

we construct a measure of digitalisation through a machine learning approach and find that the mixed ownership reform of SOEs can significantly promote digitalisation. The conclusion holds after various robustness checks, including using alternative independent and dependent variables, PSM, and endogeneity tests. The mechanism analysis reveals that mixed ownership reform facilitates digitalisation by influencing pay-performance sensitivities and risk-taking. This effect is more significant among competitive industries and corporate decision-makers who did not experience famine during childhood. Our study explores digitalisation measures, enriches the literature on the impact of mixed ownership reform on business decisions, and extends the research on the motivation of digitalisation.

The policy suggestions and managerial implications are as follows: First, it is essential to recognise the significant role of mixed ownership reform in SOEs in driving digitalisation. Promoting mixed ownership reform improves not only internal governance, enhances enterprise vitality, and increases operational efficiency but also facilitates high-quality development through the

**Table 11 Heterogeneity test: corporate decision-makers who experienced the Great Chinese Famine in childhood.**

VARIABLES	(1) Digital Famine=1	(2) Digital Famine=0	(3) Digital Famine=1	(4) Digital Famine=0
NonSOEs_Dir	0.006 (1.52)	0.015*** (4.38)		
NonSOEs_Equ			0.004 (0.02)	0.359** (2.42)
Constant	0.600 (1.29)	0.456 (0.88)	0.624 (1.35)	0.613 (1.18)
Observations	3599	6845	3599	6840
R-squared	0.394	0.333	0.392	0.325
Controls	YES	YES	YES	YES
IND FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Empirical p-value	0.000		0.000	

Note: The adjusted for clustering robust standard errors are in parentheses. Standard errors are clustered at the firm level, and significance levels are denoted as \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .

advancement of digitalisation. In implementing digital strategies for SOEs, governments need to concern institutional factors, with particular attention to the role of mixed ownership reform. Second, enterprises should leverage the influence of mixed ownership reform on pay-performance sensitivities and risk-taking attitudes. Using mixed ownership reform as an opportunity, enterprises should align their operations with value creation principles, optimise performance-based compensation mechanisms, and enhance reasonable levels of risk-taking. These measures will effectively promote both market-oriented strategies and digitalisation. Third, attention should be paid to the heterogeneous effects of mixed ownership reform on digitalisation across different types of SOEs. The influence of decision-makers' childhood experiences highlight the need for governments and enterprises to consider the personal backgrounds of executives when promoting digitalisation or technological innovation through mixed ownership reform. In competitive industries, the pace of reform can be accelerated to achieve rapid improvements in digitalisation. In non-competitive industries, modern management tools can be introduced to enhance the effect of reform on the pay-performance sensitivities of management, enabling reforms to play a greater role in driving digitalisation in non-competitive enterprises.

Our current research still has limitations: Although natural language processing (NLP) technologies have been widely applied to extract specific information from unstructured data such as text, achieving substantial progress, our measure may still have certain biases. For instance, variations in the writing style and information transparency of annual reports across different firms might lead to some firms exaggerating or obscuring their actual digitalisation efforts in their reports. This may cause the inaccuracy of the measure. In the future, we could integrate methods such as event studies to more accurately analyse and understand the role of mixed ownership reform in digitalisation.

Our research can also be further expanded: On one hand, future studies could expand to the domain of non-state-owned enterprises. To date, over 60% of privately listed firms in China have introduced state-owned shareholders through "reverse" mixed ownership reform. Further research could examine the impact of such "reverse" mixed ownership reform on corporate digitalisation. Another intriguing research direction is to extend this methodology to an international context, particularly by comparing authoritarian and democratic regimes, as well as developed and developing countries. These countries exhibit

significant differences in the role of state-owned economies, and digitalisation is influenced by factors such as policies, economic structures, and corporate cultures. Cross-national research could help us understand the similarities and differences in how reforms drive digitalisation across various economic systems.

**Data availability**

The data analysed during the current study are available in the Open Science Framework repository: <https://doi.org/10.17605/OSF.IO/W3SHV>.

Received: 14 April 2024; Accepted: 20 February 2025;

Published online: 03 March 2025

**Notes**

- 1 The abbreviation "ST" stands for "Special Treatment", which is a policy targeting stocks with abnormal operating performance. Stocks marked with "ST\*" indicate that they are at the risk of being delisted.
- 2 We take the annual report of Shanghai Modern Pharmaceutical Co., Ltd (stock code: 600420) as an example. From the resume of the board members disclosed in its 2015 annual report, we can find that director Huang Yongbing works as an assistant to the general manager of the fifth largest shareholder, Shanghai Gaodong Economic Development Co., Ltd (a non-state shareholder). Hence, we can determine that the non-state shareholder, Shanghai Gaodong Economic Development Co., Ltd, has appointed a director to participate in the governance of the listed company.
- 3 The most obvious example of this is that almost all previous studies have ignored English and specialized expressions in the context of annual reports, such as "VR", "AR", "AI", "SaaS", "PaaS", and other words, which are widely found in annual reports.

**References**

Acemoglu D, Johnson S, Robinson JA (2001) The colonial origins of comparative development: An empirical investigation. *American Economic Review* 91(5):1369–1401

Acemoglu D, Restrepo P (2018) The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment. *American Economic Review* 108(6):1488–1542. <https://doi.org/10.1257/aer.20160696>

Bardhan P (1997) Corruption and development: A review of issues. *Journal of Economic Literature* 35(3):1320–1346

Barrett M, Davidson E, Prabhu J, Vargo SL (2015) Service Innovation In The Digital Age: Key Contributions And Future Directions. *MIS Quarterly* 39(1):135–154

Bharadwaj A, Sawy O, Pavlou P, Venkatraman N (2013) Digital Business Strategy: Toward a Next Generation of Insights. *MIS Quarterly: Management Information Systems* 37:471–482. <https://doi.org/10.25300/MISQ/2013/37.2.3>

Boardman AE, Vining AR (1989) Ownership and Performance in Competitive Environments: A Comparison of the Performance of Private, Mixed, and State-Owned Enterprises. *The Journal of Law and Economics* 32(1):1–33. <https://doi.org/10.1086/467167>

Boubakri N, Cosset J-C, Saffar W (2013) The role of state and foreign owners in corporate risk-taking: Evidence from privatization. *Journal of Financial Economics* 108(3):641–658

Cai G, Liu J, Ma X (2018) Governance of Non-state Shareholders and Executive Compensation Incentives of State-owned Enterprises. *Management World* 5:137–149. <https://doi.org/10.19744/j.cnki.11-1235/f.2018.05.011>

Chen W, Srinivasan S (2024) Going digital: Implications for firm value and performance. *Review of Accounting Studies* 29(2):1619–1665

Chen X, Wan P, Ma Z, Yang Y (2024) Does corporate digital transformation restrain ESG decoupling? Evidence from China. *Humanities and Social Sciences Communications* 11(1):407. <https://doi.org/10.1057/s41599-024-02921-w>

Cheng M, Wei M (2013) The Excessive Authority of Related Shareholders. *China Industrial Economics* 307(10):108–120. <https://doi.org/10.19581/j.cnki.ciejournal.2013.10.009>

Commander S, Harrison R, Menezes-Filho N (2010) ICT and Productivity in Developing Countries: New Firm Level Evidence from Brazil and India. *The Review of Economics and Statistics* 93(2):528–541. [https://doi.org/10.1162/REST\\_a\\_00080](https://doi.org/10.1162/REST_a_00080)

Cycyota CS, Harrison D (2006) What (Not) to Expect when Surveying Executives: A Meta-Analysis of Top Manager Response Rates and Techniques Over Time. *Organizational Research Methods* 9:133–160

Dana Jr JD, Orlov E (2014) Internet penetration and capacity utilization in the US airline industry. *American Economic Journal: Microeconomics* 6(4):106–137



- Fan JP, Wong T, Zhang T (2013) Institutions and organizational structure: The case of state-owned corporate pyramids. *The Journal of Law, Economics, & Organization* 29(6):1217–1252
- Feng X, Johansson AC (2018) Living through the Great Chinese Famine: Early-life experiences and managerial decisions. *Journal of Corporate Finance* 48:638–657. <https://doi.org/10.1016/j.jcorpfin.2017.11.012>
- Firth M, Fung PMY, Rui OM (2006) Corporate performance and CEO compensation in China. *Journal of Corporate Finance* 12(4):693–714. <https://doi.org/10.1016/j.jcorpfin.2005.03.002>
- Fogel K, Morck R, Yeung B (2008) Big business stability and economic growth: Is what's good for General Motors good for America? *Journal of Financial Economics* 89(1):83–108
- Ghasemaghai M, Calic G (2019) Assessing the impact of big data on firm innovation performance: Big data is not always better data. *Journal of Business Research* 108:147–162. <https://doi.org/10.1016/j.jbusres.2019.09.062>
- Gupta N (2005) Partial privatization and firm performance. *Journal of Finance* 60(2):987–1015. <https://doi.org/10.1111/j.1540-6261.2005.00753.x>
- Hess T, Matt C, Benlian A, Wiesbock F (2016) Options for Formulating a Digital Transformation Strategy. *MIS Quarterly Executive* 15(2):123–139
- Jensen MC, Murphy KJ (1990) Performance Pay and Top-Management Incentives. *Journal of Political Economy* 98(2):225–264. <http://www.jstor.org/stable/2937665>
- Jiang F, Meng L, Tang G (2021) Media Textual Sentiment and Chinese Stock Return Predictability. *China Economic Quarterly* 21(4):1323–1344
- Jiang T (2022) Mediating Effects and Moderating Effects in Causal Inference. *China Industrial Economics* 05:100–120. <https://doi.org/10.19581/j.cnki.ciejournal.2022.05.005>
- John K, Litov L, Yeung B (2008) Corporate governance and risk-taking. *Journal of Finance* 63(4):1679–1728. <https://doi.org/10.1111/j.1540-6261.2008.01372.x>
- Kato T, Long C (2006) Executive turnover and firm performance in China. *American Economic Review* 96(2):363–367
- Ke, B, Liu, N, & Tang, S (2017). The Effect of Anti-Corruption Campaign on Shareholder Value in a Weak Institutional Environment: Evidence from China. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2963478>
- Kohli R, Johnson S (2011) Digital Transformation In Latecomer Industries: Cio And Ceo Leadership Lessons From Encana Oil & Gas (USA) Inc. *MIS Quarterly Executive* 10(4):141–156
- Korinek, A, & Stiglitz, JE (2021). Artificial intelligence, globalization, and strategies for economic development. NBER Working Paper NO.28453
- La Porta R, López-de-Silanes F, Andrei S (1999) The Benefits of Privatization: Evidence from Mexico. *The Quarterly Journal of Economics* 114(4):1193–1242. <https://doi.org/10.1162/003355399556250>
- Laffont, J.-J, & Tirole, J (1993). A Theory of Incentives in Procurement and Regulation. *The MIT Press*. <https://doi.org/10.2307/2235329>
- LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521(7553):436–444. <https://doi.org/10.1038/nature14539>
- Li K, Liu X, Mai F, Zhang T (2021a) The Role of Corporate Culture in Bad Times: Evidence from the COVID-19 Pandemic. *Journal of Financial and Quantitative Analysis* 56:1–68. <https://doi.org/10.1017/S0022109021000326>
- Li K, Mai F, Shen R, Yan XY (2021b) Measuring Corporate Culture Using Machine Learning. *Review of Financial Studies* 34(7):3265–3315. <https://doi.org/10.1093/rfs/hhaz079>
- Li L, Su F, Zhang W, Mao JY (2018) Digital transformation by SME entrepreneurs: A capability perspective. *Information Systems Journal* 28(6):1129–1157. <https://doi.org/10.1111/isj.12153>
- Li W, Yu M (2015) Equity structure of privatized enterprises and innovation. *Management World* 4:112–125. <https://doi.org/10.19744/j.cnki.11-1235/f.2015.04.011>
- Lin C, Lin P, Song FM, Li C (2011) Managerial incentives, CEO characteristics and corporate innovation in China's private sector. *Journal of Comparative Economics* 39(2):176–190
- Lin JY, Cai F, Li Z (1998) Competition, Policy Burdens, and State-Owned Enterprise Reform. *The American Economic Review* 88(2):422–427. <http://www.jstor.org/stable/116960>
- Loebbecke C, Picot A (2015) Reflections on societal and business model transformation arising from digitization and big data analytics: A research agenda. *Journal of Strategic Information Systems* 24(3):149–157. <https://doi.org/10.1016/j.jsis.2015.08.002>
- Loughran T, McDonald B (2011) When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *Journal of Finance* 66(1):35–65. <https://doi.org/10.1111/j.1540-6261.2010.01625.x>
- Marabelli M, Galliers RD (2017) A reflection on information systems strategizing: the role of power and everyday practices. *Information Systems Journal* 27(3):347–366. <https://doi.org/10.1111/isj.12110>
- Matt C, Hess T, Benlian A (2015) Digital Transformation Strategies. *Business & Information Systems Engineering* 57(5):339–343. <https://doi.org/10.1007/s12599-015-0401-5>
- Meggison WL, Nash RC, Van Randenborgh M (1994) The financial and operating performance of newly privatized firms: An international empirical analysis. *The Journal of Finance* 49(2):403–452
- Meggison WL, Netter JM (2001) From state to market: A survey of empirical studies on privatization. *Journal of Economic Literature* 39(2):321–389
- Müller O, Fay M, Brocke JV (2018) The Effect of Big Data and Analytics on Firm Performance: An Econometric Analysis Considering Industry Characteristics. *Journal of Management Information Systems* 35:488–509. <https://doi.org/10.1080/07421222.2018.1451955/>
- Opoku-Mensah E, Yin Y, Obiora SC, Adjei Darko P (2022) Grabbing hand or helping hand? Ownership interventions and acquirers returns; the role of provincial idiosyncrasies. *International Journal of Finance & Economics* 27(4):4797–4815. <https://doi.org/10.1002/ijfe.2401>
- Peng Y, Wang N, Gu L (2023) Enterprise digital transformation, predictive information disclosure and the stock price crash risk. *Finance & Trade Economics* 44(5):73–90
- Porter ME, Heppelmann JE (2014) How Smart, Connected Products Are Transforming Competition. *Harvard Business Review* 92(11):64–88
- Rosenblat A, Stark L (2016) Algorithmic Labor and Information Asymmetries: A Case Study of Uber's Drivers. *International Journal of Communication* 10:3758–3784
- Scholz RW, Czichos R, Parycek P, Lampoltshammer TJ (2020) Organizational vulnerability of digital threats: A first validation of an assessment method. *European Journal of Operational Research* 282(2):627–643. <https://doi.org/10.1016/j.ejor.2019.09.020>
- Shleifer A, Vishny R (1994) Politicians and Firms. *The Quarterly Journal of Economics* 109:995–1025. <https://doi.org/10.2307/2118354>
- Singh A, Hess T (2017) How Chief Digital Officers Promote the Digital Transformation of their Companies. *MIS Quarterly Executive* 16:1–17
- Svahn F, Mathiassen L, Lindgren R (2017) Embracing Digital Innovation In Incumbent Firms: How Volvo Cars Managed Competing Concerns. *MIS Quarterly* 41(1):239–253. <https://doi.org/10.25300/misq/2017/41.1.12>
- Triplet J (1999) The Solow Productivity Paradox: What do Computers do to Productivity? *Canadian Journal of Economics* 32:309–334. <https://doi.org/10.2307/136425>
- Verhoef P, Broekhuizen T, Bart Y, Bhattacharya A, Dong J, Fabian N, Haenlein M (2019) Digital transformation: A multidisciplinary reflection and research agenda. *Journal of Business Research* 122:889–901. <https://doi.org/10.1016/j.jbusres.2019.09.022>
- Vial G (2019) Understanding digital transformation: A review and a research agenda. *The Journal of Strategic Information Systems* 28(2):118–144. <https://doi.org/10.1016/j.jsis.2019.01.003>
- Wang, Z, Zhou, S, Zhou, Y, & Yang, H (2022). The impact of CEOs' early-life experience on the financialization of non-financial firms: Evidence from the great Chinese famine. *International Journal of Finance & Economics*, 1–18. <https://doi.org/10.1002/ijfe.2755>
- Warner KSR, Wäger M (2019) Building dynamic capabilities for digital transformation: An ongoing process of strategic renewal. *Long Range Planning* 52(3):326–349. <https://doi.org/10.1016/j.lrp.2018.12.001>
- Wu F, Hu H, Lin H, Ren X (2021) Enterprise Digital Transformation and Capital Market Performance: Empirical Evidence from Stock Liquidity. *Management World* 37(07):130–144+110. <https://doi.org/10.19744/j.cnki.11-1235/f.2021.0097>
- Wu Y, Li Z (2024) Digital transformation, entrepreneurship, and disruptive innovation: evidence of corporate digitalization in China from 2010 to 2021. *Humanities and Social Sciences Communications* 11(1):163. <https://doi.org/10.1057/s41599-023-02378-3>
- Xu N, Li Z (2016) CEOs' Poverty Experience and Corporate Philanthropy. *Economic Research Journal* 51(12):133–146
- Yuan C, Xiao T, Geng C, Sheng Y (2021) Digital Transformation and Division of Labor between Enterprises: Vertical Specialization or Vertical Integration. *China Industrial Economics* 9:137–155. <https://doi.org/10.19581/j.cnki.ciejournal.2021.09.007>
- Zhao C, Wang W, Li X (2021) How Does Digital Transformation Affect the Total Factor Productivity of Enterprises? *Finance & Trade Economics* 42(7):114–129. <https://doi.org/10.109795/j.cnki.cn11-1166/f.20210705.001>

## Acknowledgements

This study is supported by the “2024-2025 Guizhou Provincial Youth Development Research Project” (Grant: QNYB24110).

## Author contributions

WT was responsible for the research design, data collection, and preliminary analysis. He also participated in the writing and revision process of the manuscript, providing key insights into the theoretical framework and empirical analysis. W-CH played a significant role in data preprocessing and the establishment of statistical models. NM was primarily responsible for the literature review section, systematically collating research progress in the relevant field and providing a solid theoretical foundation for the paper. JH provided valuable suggestions and feedback during the writing process. She was responsible for interpreting the research results and ensuring that the conclusions of the paper were clear and reasonable. She also participated in the proofreading and formatting review of the manuscript to enhance its overall quality.



**Competing interests**

The authors declare no competing interests.

**Ethical approval**

Ethical approval was not required as the study did not involve human participants.

**Informed consent**

This article does not contain any studies with human participants performed by any of the authors.

**Additional information**

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1057/s41599-025-04613-5>.

**Correspondence** and requests for materials should be addressed to Juan He.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025