MASTER
CIÊNCIAS EMPRESARIAIS

MASTER´S FINAL WORK
DISSERTATION

DIGITAL TRANSFORMATION AND FIRM PERFORMANCE: HETEROGENEOUS EVIDENCE FROM CHINA AND THE EUROPEAN UNION UNDER THE COVID-19 SHOCK

XIAOXIN LI

JUNE-2026

MASTER
CIÊNCIAS EMPRESARIAIS

MASTER´S FINAL WORK
DISSERTATION

DIGITAL TRANSFORMATION AND FIRM PERFORMANCE: HETEROGENEOUS EVIDENCE FROM CHINA AND THE EUROPEAN UNION UNDER THE COVID-19 SHOCK

XIAOXIN LI

SUPERVISION:
PROFESSOR DOUTOR JOÃO MANUEL JORGE ESTEVÃO

JUNE-2026

Abstract

This study examines the relationship between digital transformation and firm performance using a balanced panel of listed companies in China and the European Union from 2015 to 2024, comprising 2,512 firms and 25,120 firm-year observations. Digital transformation is examined along two dimensions: text-based strategic attention to digitalization and asset-output efficiency as an indirect proxy for operational conversion. The paper further examines whether the relationship between digital transformation and firm performance varies across institutional environments and under the COVID-19 context.

The results from firm and year fixed effects models show that digital transformation does not have a single and stable positive effect on firm performance. Instead, its effects vary across digitalization dimensions, performance indicators, and time horizons. Text-based digital attention is positively associated with accounting performance, but its lagged effect becomes negative, suggesting that strategic attention to digitalization does not necessarily translate into subsequent performance improvements in the short term. In contrast, the indicator based on asset output efficiency shows stronger explanatory power in the lagged models. Further analysis shows stronger marginal associations in the selected EU sample, while asset-output efficiency is more closely associated with pandemic-period accounting performance than textual digital attention.

The originality of this study lies in reinterpreting digital transformation through the dimensions of strategic attention and operational conversion, while integrating China-EU comparison and the COVID-19 shock into the same analytical framework. The findings show that textual digital attention and asset-output efficiency represent distinct dimensions of digital transformation, with different performance associations across time, institutional contexts, and crisis conditions.

Keywords: Digital transformation; Firm performance; Textual analysis; Operational efficiency; Institutional heterogeneity

Resumo

Este estudo analisa a relação entre transformação digital e desempenho empresarial, com base num painel balanceado de empresas cotadas na China e na União Europeia entre 2015 e 2024, composto por 2.512 empresas e 25.120 observações empresa-ano. A transformação digital é examinada em duas dimensões: a atenção estratégica à digitalização baseada em informação textual e a eficiência de produção dos ativos como proxy indireta da conversão operacional. O estudo examina ainda se a relação entre transformação digital e desempenho empresarial varia entre contextos institucionais e no contexto da COVID-19.

Os resultados dos modelos de efeitos fixos de empresa e de ano mostram que a transformação digital não apresenta um efeito positivo único e estável sobre o desempenho. Os efeitos variam consoante a dimensão da digitalização, o indicador de desempenho e o horizonte temporal. A atenção digital baseada em texto associa-se positivamente ao desempenho contabilístico, mas o seu efeito desfasado torna-se negativo, sugerindo que a atenção estratégica à digitalização não se converte necessariamente em melhorias de desempenho no curto prazo. Em contraste, o indicador baseado na eficiência de produção dos ativos apresenta maior poder explicativo nos modelos desfasados. A análise adicional mostra associações marginais mais fortes na amostra selecionada da UE, enquanto a eficiência de produção dos ativos está mais associada ao desempenho contabilístico no contexto pandémico do que a atenção digital textual.

A originalidade deste estudo reside em reinterpretar a transformação digital através das dimensões de atenção estratégica e conversão operacional, integrando a comparação China-União Europeia e o choque da COVID-19 no mesmo quadro analítico. Os resultados mostram que a atenção digital textual e a eficiência de produção dos ativos representam dimensões distintas da transformação digital, com diferentes associações com o desempenho ao longo do tempo e em contextos institucionais e de crise.

Palavras-chave: Transformação digital; Desempenho empresarial; Análise textual; Eficiência operacional; Heterogeneidade institucional

Acknowledgments

At the moment of completing this thesis, I feel as if I am slowly waking from a long and demanding journey. Looking back on my time in Portugal, I remember the curiosity I had when I first arrived, as well as the uncertainty, pressure, and loneliness of living in a different language and cultural environment. Yet I also know that I never faced these challenges alone, and it is with genuine gratitude that I write these words.

First and foremost, I would like to express my sincere gratitude to my supervisor, Professor João Manuel Jorge Estevão, for his patient guidance, professional advice, and continuous encouragement throughout this thesis. From refining the research direction and model specification to improving the thesis structure and interpreting the empirical results, his suggestions helped me clarify my ideas. At many moments when I felt lost, each meeting helped me regain direction and confidence. I am especially grateful for his tolerance and patience with my language expression and research difficulties.

Under his guidance, writing this thesis became not only an academic task, but also a process of learning how to face uncertainty. Facing the difficulties that arose during this journey, I gradually learned to break down complex problems, move forward step by step, and trust myself through patience and sustained effort.

I would also like to thank the classmates and friends I met during my studies in Portugal. I first arrived in a foreign country, many things felt unfamiliar and uneasy, but Mariana, Inês, Margarida, João, and many kind classmates offered me kindness and support. It was also a true gift to know Beichen Xu and Bixiao Sun. From celebrating festivals and birthdays together to travelling through different countries, they filled my life with warmth, laughter, and a sense of belonging. These shared moments will remain among the most precious memories of this journey.

Finally, I would like to thank my parents for their constant financial and emotional support, for respecting my choices, and for believing in me even when I doubted myself. I also thank my forever friends Ling Yang and Ziqing Liu for their companionship and encouragement. Lastly, I want to thank my little cat, who quietly stayed beside my computer on many writing days and brought me comfort.

Index

List of Tables

List of Abbreviations

CSMAR - China Stock Market and Accounting Research

DCV - Dynamic Capability View

EU - European Union

GDPR - General Data Protection Regulation

OxCGRT - Oxford COVID-19 Government Response Tracker

RBV - Resource-based View

ROA - Return on Assets

ROD - Return on Digitalization

ROE - Return on Equity

VIF - Variance Inflation Factor

1. Introduction

In today's global economy, digital transformation has become one of the key strategic changes facing firms across industries (Li *et al.*, 2025). Digital technology is widely regarded as an important source of competitive advantage, helping firms improve operational efficiency and respond more flexibly to changes in the external environment (Barney, 1991; Teece *et al.*, 1997; Zeng *et al.*, 2022). A growing body of empirical research shows that firms with higher levels of digitalization tend to exhibit better financial performance and higher market valuations, as digital transformation helps improve resource allocation, drive business-model innovation, and enhance firms' information-processing capacity (Mahssouni *et al.*, 2023; Mao & Wang, 2024; Zhou *et al.*, 2023). At the same time, existing studies also note that the relationship between digital transformation and firm performance is not always consistently positive. The "digitalization paradox" suggests that high upfront costs, restructuring pressures, and implementation lags may suppress short-term profitability during long-term capability formation (Gebauer *et al.*, 2020; Guo & Xu, 2021). Digital transformation is therefore both strategically significant and, empirically, a relatively complex issue.

However, regarding when and how digital transformation translates into improved firm performance, the literature still has several research gaps. First, most studies treat digital transformation as a homogeneous concept, with limited distinction between the strategic attention reflected in textual disclosure and firms' capability to convert resource inputs into operating output (Axenbeck & Breithaupt, 2022; Yavuz *et al.*, 2025). Second, comparative research across institutional environments remains limited. China and the EU differ markedly in digital-governance logic and in the state-market relationship, which may shape how digital transformation translates into firm performance (Cai & Zhang, 2022; Zheng & Snyder, 2023). Third, as a major external shock, the COVID-19 pandemic accelerated the adoption of remote work, online services, and digital collaboration, and may also have altered the relative importance of different digital-transformation dimensions for firm outcomes (Bai *et al.*, 2021; Chen *et al.*, 2024).

To address these gaps, this paper draws on firm-level panel data covering listed companies in China and the EU from 2015 to 2024. Two complementary digital-transformation indicators are constructed. The first is the text-based indicator lnTextDigital, built from keyword analysis of annual reports, which measures a firm's strategic attention to digitalization issues and the intensity of its disclosure. The second is ROD2, which measures a firm's asset output efficiency as the ratio of operating revenue to total assets, and serves as an indirect proxy for operational conversion capacity that may be related to digital transformation. Firm performance is measured by Return on Assets (ROA) and Return on Equity (ROE) for accounting performance and by Tobin's Q for market valuation. This paper employs panel regression models to estimate, respectively, the baseline effects, the lagged effects, and heterogeneous effects across different institutional environments and pandemic periods. This design allows the paper to examine not only the relationship between digital transformation and firm performance, but also how this relationship varies across institutional environments and pandemic periods.

This paper contributes to the existing literature in three ways. First, this study develops a dual-path analytical framework for digital transformation by distinguishing between strategic attention in textual disclosure and operational conversion reflected in asset output efficiency, thereby clarifying how different digitalization-related indicators are associated with firm performance. Second, by comparing listed companies in China and the EU within a unified empirical framework, this paper examines how the association between digitalization indicators and firm performance differs across the China and selected EU samples. Third, this paper incorporates the COVID-19 pandemic as a contextual factor, examining whether the associations between different digitalization dimensions and firm performance vary under external crisis conditions. Overall, this paper suggests that digital transformation should be understood as a process that varies across contexts, with its performance associations shaped by how digitalization is expressed, how it is reflected in operational conversion, the institutional environment in which the firm operates, and the external conditions it faces.

The remainder of this paper is organized as follows. Section 2 develops the theoretical framework and the research hypotheses. Section 3 presents the data and the empirical methodology, covering the sample, the variable construction, and the model specifications. Section 4 reports the empirical results. Section 5 discusses the findings and their theoretical and practical implications. Section 6 concludes the paper and outlines its limitations.

2. Literature Review

2.1 Strategic Digital Attention and Operational Conversion

Digital transformation is generally understood as the process by which firms embed digital technologies into business processes, organizational structures, and modes of value creation, a concept that extends beyond data-format conversion or the adoption of digital technology (Gradillas & Thomas, 2023; Liu *et al.*, 2011; Rachinger *et al.*, 2019). At the firm level, this process manifests along multiple dimensions, and existing research has mainly identified two observable paths. The textual-disclosure path constructs a digitalization index based on the frequency of keywords such as big data, artificial intelligence, and cloud computing appearing in annual reports, reflecting management's strategic attention and disclosure intent (Zeng *et al.*, 2022; Mao & Wang, 2024). However, this approach is highly dependent on dictionary construction and contextual judgment, as extensive references to digitalization in annual reports may reflect genuine strategic investment or amount to symbolic disclosure not yet validated by operating outcomes, with its credibility resting on the constraints imposed by external verification mechanisms (Loughran & McDonald, 2016; Healy & Palepu, 2001). The operational conversion path, by contrast, uses efficiency indicators to observe firms' resource utilization and operating output (Yavuz *et al.*, 2025). In this study, this path is proxied indirectly by ROD2, which captures asset output efficiency rather than digitalization itself. These two indicators respectively capture the dimensions of strategic disclosure and operational conversion, forming the basis of this paper's dual-path analytical framework.

From a theoretical perspective, these two paths follow different performance-conversion logics. Resource-based theory (RBV) holds that data assets, algorithmic capability, and digital technology embedded in business processes are able to confer sustained competitive advantage on firms because these resources are valuable, scarce, and difficult to imitate, making it hard for competitors to obtain the same advantage through simple replication (Barney, 1991; Wade & Hulland, 2004). Dynamic capabilities theory (DCV) further suggests that, in highly uncertain environments, firms must rely on three dynamic capabilities, namely sensing opportunities, seizing opportunities, and reconfiguring resources, to adapt to change, and whether digital transformation can effectively improve performance depends on whether it genuinely strengthens a firm's capacity for resource reconfiguration and environmental adaptation (Teece *et al.*, 1997; Helfat & Winter, 2011). Accordingly, the competitive advantage of digital transformation depends on the actual deployment of digital resources and the genuine accumulation of dynamic capabilities, rather than on digitalization expressed merely at the strategic level.

Most empirical studies show that digital transformation exerts a significant positive effect on financial performance and market value, a conclusion widely validated across samples of European Union (EU) listed companies, small and medium-sized enterprises, and Chinese firms (Mahssouni *et al.*, 2023; Zeng *et al.*, 2022; Kádárová *et al.*, 2023). However, different performance indicators capture different economic meanings: accounting-performance indicators such as ROA and ROE mainly reflect contemporaneous profitability, whereas market-value indicators such as Tobin's Q more strongly reflect capital-market expectations regarding a firm's future growth potential (Al-Matari *et al.*, 2014). Chen and Srinivasan (2024) find that, after non-technology firms disclose digitalization activities, market valuation improves and asset turnover increases, but ROA does not improve immediately, indicating that digitalization affects different performance dimensions through different paths.

The effect of digital transformation on firm performance is realized gradually. There is a clear time lag between IT investment and productivity improvement, and firms typically must go through complementary investment processes, such as business-process re-engineering, organizational restructuring, and the accumulation of employee skills, before the economic effects of technology can be progressively released (Brynjolfsson, 1993; Brynjolfsson & Hitt, 2000). This phenomenon is generally summarized as the "digital paradox," whereby high upfront investment during the early stage of transformation may suppress short-term profitability (Gebauer *et al.*, 2020; Zhou *et al.*, 2023). Guo and Xu (2021) find that the performance effects of digital transformation in Chinese manufacturing firms follow a typical U-shaped trajectory. Jardak & Ben Hamad (2022) likewise find, in a Swedish sample, that digitalization places downward pressure on profits in the short term while having a positive effect on market valuation. Chen and Srinivasan (2024) further show that capital markets may respond to firms' digitalization signals relatively early, whereas improvements in accounting performance typically take longer to materialize.

The evidence above suggests that the relationship between digital transformation and firm performance differs substantially across digitalization dimensions. The performance associations of strategic digital attention reflected in textual disclosure may be influenced by the fulfillment of strategic commitments, organizational adjustment costs, and disclosure credibility, and is therefore subject to greater uncertainty and time lag. By contrast, ROD2 provides an indirect observation window for asset-output efficiency and operational conversion capacity, rather than a direct measure of digital transformation itself, and may therefore show a more stable association with firm performance. Based on this reasoning, this paper proposes the following hypothesis:

H1. ROD2 is expected to show a more stable positive association with firm performance over time than lnTextDigital.

2.2 Institutional Heterogeneity: China and the EU

The institutional environment constitutes one of the important contexts for understanding the performance effects of digital transformation (Wang & Han, 2026). China and the EU differ significantly along three dimensions: data governance, digital industrial strategy, and market-regulatory tools (Cai & Zhang, 2022; Wang & Han, 2026; Zheng & Snyder, 2023). In terms of data governance, the EU centers on personal-data protection and the construction of unified rules, with the General Data Protection Regulation (GDPR) advancing the protection of data-subject rights and the coordination of rules across member states (Liang & Wu, 2020). China, by contrast, places greater emphasis on utilizing data as a factor of production while safeguarding data security and national sovereignty, with its institutional logic placing greater emphasis on security orientation and state regulation (Nanni *et al.*, 2024). In terms of digital industrial strategy, the EU shapes the institutional boundaries of the digital market through rule-setting and standard-setting (Cai & Zhang, 2022), whereas China places greater emphasis on a developmental industrial-policy logic, relying on national-level industrial planning to guide firms in advancing technological upgrading (Li, 2017). In terms of market regulation, the EU tends to establish market rules through legally binding hard law, whereas China relies more on flexible policy guidance to address issues of digital-competition governance (Zheng & Snyder, 2023).

These differences suggest that Chinese and EU firms operate in different institutional environments, which may shape the observed associations between digitalization indicators and firm performance. The EU's more formalized digital governance framework may provide clearer institutional expectations regarding information disclosure, data use, and market competition. This may be associated with stronger marginal relationships between digitalization indicators and performance in the selected EU sample. In contrast, China's digital transformation process is more strongly shaped by policy promotion and rapid diffusion, which may make firm level marginal differences less stable. Accordingly, the following hypothesis is proposed:

H2. The associations of lnTextDigital and ROD2 with firm performance are expected to be stronger in the selected EU sample than in the China sample.

2.3 COVID-19 Shock and Operational Resilience

The sudden outbreak of COVID-19 triggered a severe global public-health crisis. Pandemic containment measures restricted the movement of people, disrupted offline operations, and interrupted global supply chains, exposing firms to multiple pressures including a sharp drop in demand, rising operating costs, and tighter financing constraints (Bai *et al.*, 2021; Ding *et al.*, 2021; Wu *et al.*, 2023). Under these conditions, firms were forced to rapidly adapt to the new normal of remote work and online operations (Herath & Herath, 2020).

Existing research generally holds that digital capability is a key factor enabling firms to mitigate external shocks and sustain ongoing operations (Härting *et al.*, 2022). Digital technology supports remote work, online sales, and digital collaboration, effectively reducing the coordination costs arising from physical isolation (Brynjolfsson *et al.*, 2020; Ding *et al.*, 2021). And firms with higher levels of digitalization generally exhibit stronger operational resilience and performance during crises (Chen *et al.*, 2024; Huang & Zhao, 2024). In addition, digital transformation may strengthen firms' operational resilience by improving inventory turnover and capital use efficiency, reducing supply chain risk through more diverse customer and supplier networks, and easing liquidity pressure through better access to external financing, especially among small and medium enterprises (Chen *et al.*, 2024; Guo *et al.*, 2020; Wu *et al.*, 2021).

However, without corresponding adjustments to organizational structure, management processes, and corporate culture, digitalization adopted as an emergency response to the pandemic may increase organizational burden and undermine performance over the long term (Reuschl *et al.*, 2022). In a crisis context, market participants and external financiers may adopt a more cautious stance toward strategic commitments and textual disclosure signals that have not yet been supported by operating outcomes, while firms that can effectively convert assets and resources into operating output are better positioned to buffer the pandemic shock and sustain performance (Chen *et al.*, 2024; Reuschl *et al.*, 2022). The performance effect of digital transformation indicators during the pandemic may therefore depend more on whether firms can embed digital technology into operational coordination and resource allocation processes than on strategic level signaling alone. Accordingly, the following hypothesis is proposed:

H3. Under the COVID-19 shock and stricter policy constraints, ROD2 is expected to show a stronger positive association with firm performance than lnTextDigital.

3. Data and Methodology

3.1 Sample and Data

This paper uses listed companies in mainland China and the EU from 2015 to 2024 as the research sample. To maintain consistency between the sample scope and regional data matching, the China sample is restricted to mainland Chinese listed companies, excluding firms from Hong Kong, Macao, and Taiwan. Financial data for Chinese listed firms are obtained from the China Stock Market and Accounting Research (CSMAR) database, while financial data for EU listed firms are obtained from the Bloomberg database. Annual reports for firms in both regions were obtained from their respective corporate official websites.

The sample selection process follows two principles. First, firms with missing key variables or incomplete annual reports during the study period are excluded to ensure the continuity of the panel structure (Hsiao, 2007). Second, for text matching, Chinese listed firms' annual reports are used directly in Chinese, while EU listed firms' annual reports are used consistently in English. This approach maintains a consistent text matching procedure and reduces measurement bias arising from multilingual disclosure within the EU sample. (Loughran & McDonald, 2011). After screening, the EU sample was reduced from an original 5,418 firms to 587 firms, and the China sample was reduced from 5,720 firms to 1,925 firms. In the end, this paper constructs a balanced panel covering 2,512 listed companies and 25,120 firm-year observations. Since the EU sample consists of listed firms with complete English annual reports and complete panel data during the study period, the findings related to this sample mainly apply to this group of firms.

For the pandemic data, this paper uses annual regional confirmed COVID-19 case data for 2020-2022. Data for China are obtained from the CSMAR database and matched at the provincial level, while data for the EU are obtained from the WHO database and matched at the country level. Population data are obtained respectively from the National Bureau of Statistics of China and the World Bank, and are used to construct the COVID_Shock indicator. Policy stringency is measured using the Stringency Index published by The Oxford Covid-19 Government Response Tracker (OxCGRT), which ranges from 0 to 100, with higher values indicating stricter containment policies (Hale *et al.*, 2021).

Python was used for annual-report text processing, keyword matching, and keyword-frequency calculation. Stata was used for data cleaning and merging, descriptive statistical analysis, panel regression estimation, and robustness checks.

3.2 Variables

This study constructs variables in four groups: digital transformation indicators, firm performance indicators, contextual variables, and control variables. Detailed definitions, measurement methods, and data sources are reported in Table 1.

3.2.1 Digital Transformation Indicators

This paper captures firms' digital transformation characteristics along two dimensions: textual disclosure and operational conversion. The textual disclosure dimension is measured by lnTextDigital, constructed as the natural logarithm of one plus the frequency of digitalization related keywords in annual reports. The keyword dictionary is based on Zhao *et al.* (2021) and covers major digital technologies such as big data, cloud computing, artificial intelligence, the Internet of Things, and blockchain. For EU firms' English annual reports, the Chinese keyword dictionary is translated and verified through a controlled procedure, with the complete dictionary reported in Appendix Table 1. Examples from the sentence level contextual review of the English keyword matching procedure are provided in Appendix Table 2.

The operational conversion dimension is measured by ROD2, calculated as operating revenue divided by total assets. ROD2 captures asset output efficiency and is used as an indirect operational conversion indicator rather than a direct measure of digital transformation. The original ROD indicator proposed by Yavuz *et al.* (2025), based on operating revenue per employee, is used as an alternative indicator in the robustness checks.

3.2.2 Firm Performance

Firm performance is measured from two perspectives: accounting performance and market valuation. Accounting performance is captured by ROA and ROE, while market valuation is captured by Tobin's Q. This combination allows the analysis to distinguish between short-term profitability and market expectations regarding firm value.

3.2.3 Contextual Variables

To examine institutional heterogeneity, this paper defines the EU dummy variable, which equals 1 for EU firms and 0 for Chinese firms. To capture the pandemic context, this paper defines During for 2020-2022 and Post for 2023-2024, and constructs interaction terms between these stage variables and the two digital transformation indicators. In addition, COVID_Shock measures pandemic transmission intensity, while Stringency_during captures the intensity of containment policies during 2020-2022.

3.2.4 Control Variables

The models include firm size, leverage, cash flow, fixed asset intensity, and the price-to-book ratio as control variables. These variables control for firm scale, financial structure, internal financing capacity, asset structure, and growth opportunities. The price-to-book ratio is excluded from models using Tobin's Q as the dependent variable because both variables reflect market valuation. Definitions of all variables are presented in Table 1.

Table 1: Description of Variables

Variable Type	Name	Description	Authors
Dependent Variable	ROA	ROA is measured as net profit divided by total assets and expressed in percentage points. It reflects firm profitability from an accounting perspective.	Guo & Xu (2021); Mahssouni *et al.* (2023); Yavuz *et al.* (2025)
	ROE	ROE is calculated as net income divided by shareholders' equity and expressed in percentage points. It reflects the return generated for shareholders.	Mahssouni *et al.* (2023); Gu *et al.* (2025)
	Tobin's Q	Tobin's Q is defined as the ratio of a firm's market value to the replacement cost of its assets, reflecting investor expectations and firm value	Yavuz *et al.* (2025); Mahssouni *et al.* (2023)
Independent Variable	lnTextDigital	lnTextDigital is measured as the natural logarithm of one plus the frequency of digitalization-related keywords in annual reports. It captures firms' strategic attention to digital transformation in textual disclosure.	Mao & Wang (2024); Yavuz *et al.* (2025); Zeng *et al.* (2022)
	ROD2	ROD2 is calculated as revenue divided by total assets and is used to capture a firm's asset output efficiency and resource utilization capacity. It provides a complementary operational indicator that may be related to digital transformation.	Adapted by the author from Yavuz *et al.* (2025)
Contextual variable	EU	Regional dummy variable equal to 1 for EU firms and 0 for Chinese firms; used to construct interaction terms for examining institutional heterogeneity between China and the EU.	Author
	COVID Shock	COVID Shock is measured as confirmed COVID-19 cases divided by total population during 2020-2022 and reflects the severity of the pandemic shock	Huang & Zhao (2024)
	Stringency Index	Measures the strictness of government COVID-19 containment and closure policies, based on the OxCGRT database. The index ranges from 0 to 100, with higher values indicating stricter policy responses.	Hale *et al.* (2021)

Control Variable	Size	Firm size is measured as the natural logarithm of the number of employees	Gu *et al.* (2025); Mahssouni *et al.* (2023)
	Leverage	Leverage is measured as total liabilities divided by total assets and reflects the firm's financial structure	Mahssouni *et al.* (2023)
	Cash Flow	Cash Flow is measured as operating cash flow divided by total assets and reflects internal financing capacity. It is also used as an alternative dependent variable in the robustness test.	Huang & Zhao (2024); Mahssouni *et al.* (2023)
	Fixed Asset Intensity	Fixed Asset Intensity is measured as fixed assets divided by total assets and reflects the firm's asset structure	Huang & Zhao (2024); Mao & Wang (2024)
	Price-to-Book Ratio	Price-to-Book Ratio is calculated as market value divided by book value and proxies firm growth opportunities and market expectations	Huang & Zhao (2024); Mahssouni *et al.* (2023)

Source: Author's elaboration

3.3 Empirical Model

3.3.1 Panel Model Specification

This paper employs a two-way fixed effects model incorporating firm fixed effects (μ_i) and year fixed effects (λ_t) to control for unobserved, time-invariant firm heterogeneity and common year shocks (Wooldridge, 2009). All models use robust standard errors clustered at the firm level (Cameron & Miller, 2015). The use of a balanced panel helps ensure temporal consistency for comparisons across regions, but the conclusions of this study apply mainly to firms that survived continuously and maintained complete disclosure throughout the study period. Extreme observations in the original sample are retained in the main regressions, while a separate 1% winsorization is applied in the robustness checks.

3.3.2 Baseline Model

This paper separates the core explanatory variables into the text based digitalization indicator lnTextDigital and the asset output efficiency indicator ROD2, and constructs the following baseline models, respectively:

$$Performance_{i,t} = \beta_0 + \beta_1 \ln TextDigital_{i,t} + \gamma Controls_{i,t} + \mu_i + \lambda_t + \varepsilon_{i,t} \qquad (1)$$

$$Performance_{i,t} = \beta_0 + \beta_1 ROD2_{i,t} + \gamma Controls_{i,t} + \mu_i + \lambda_t + \varepsilon_{i,t} \qquad (2)$$

where $Performance_{i,t}$ is measured by ROA, ROE, and Tobin's Q. $Controls_{i,t}$ is the vector of control variables; μ_i and λ_t are firm and year fixed effects, respectively, and $\varepsilon_{i,t}$ is the error term.

3.3.3 Cross-Regional Heterogeneity Model

To examine differences in the digitalization effect across the institutional environments of China and the EU, this paper introduces interaction terms between the digitalization variables and the EU dummy variable, where β_2 captures the differential effect of the EU relative to China:

$$\begin{aligned} Performance_{i,t} &= \beta_0 + \beta_1 \ln TextDigital_{i,t} + \beta_2 (\ln TextDigital_{i,t} \times EU_i) \\ &+ \gamma Controls_{i,t} + \mu_i + \lambda_t + \varepsilon_{i,t} \end{aligned} \tag{3}$$

$$\begin{aligned} Performance_{i,t} &= \beta_0 + \beta_1 ROD2_{i,t} + \beta_2 \left(ROD2_{i,t} \times EU_i\right) + \gamma Controls_{i,t} + \mu_i \\ &+ \lambda_t + \varepsilon_{i,t} \end{aligned} \tag{4}$$

3.3.4 Pandemic Scenario Extension Model

To examine whether the pandemic context altered the relationship between digital transformation indicators and firm performance, this paper constructs extended models from three perspectives: pandemic stage, policy constraint intensity, and pandemic transmission intensity.

In terms of pandemic stages, this paper defines 2020-2022 as the pandemic period and 2023-2024 as the post-pandemic stage, and constructs interaction terms between the digital transformation indicators and the During and Post variables:

$$\begin{aligned} Performance_{i,t} &= \beta_0 + \beta_1 \ln TextDigital_{i,t} + \beta_2 (\ln TextDigital_{i,t} \times During_t) \\ &+ \beta_3 (\ln TextDigital_{i,t} \times Post_t) + \gamma Controls_{i,t} + \mu_i + \lambda_t + \varepsilon_{i,t} \end{aligned} \tag{5}$$

$$\begin{aligned} Performance_{i,t} &= \beta_0 + \beta_1 ROD2_{i,t} + \beta_2 (ROD2_{i,t} \times During_t) \\ &+ \beta_3 (ROD2_{i,t} \times Post_t) + \gamma Controls_{i,t} + \mu_i + \lambda_t + \varepsilon_{i,t} \end{aligned} \tag{6}$$

For policy constraint intensity, this paper constructs interaction terms between $Stringency\_during_{r,t}$ and the digital transformation indicators, as follows:

$$Performance_{i,t} = \beta_0 + \beta_1 \ln TextDigital_{i,t} + \beta_2 Stringency\_during_{r,t} + \beta_3(\ln TextDigital_{i,t} \times Stringency_{during_{r,t}}) + \gamma Controls_{i,t} + \mu_i + \lambda_t + \varepsilon_{i,t} \tag{7}$$

$$Performance_{i,t} = \beta_0 + \beta_1 ROD2_{i,t} + \beta_2 Stringency\_during_{r,t} + \beta_3(ROD2_{i,t} \times Stringency\_during_{r,t}) + \gamma Controls_{i,t} + \mu_i + \lambda_t + \varepsilon_{i,t} \tag{8}$$

To examine the role of pandemic transmission intensity, this paper further uses $COVIDShock\_during_{r,t}$. The corresponding model is specified as follows:

$$Performance_{i,t} = \beta_0 + \beta_1 \ln TextDigital_{i,t} + \beta_2 COVIDShock\_during_{r,t} + \beta_3(\ln TextDigital_{i,t} \times COVIDShock_{during_{r,t}}) + \gamma Controls_{i,t} + \mu_i + \lambda_t + \varepsilon_{i,t} \tag{9}$$

$$Performance_{i,t} = \beta_0 + \beta_1 ROD2_{i,t} + \beta_2 COVIDShock\_during_{r,t} + \beta_3(ROD2_{i,t} \times COVIDShock\_during_{r,t}) + \gamma Controls_{i,t} + \mu_i + \lambda_t + \varepsilon_{i,t} \tag{10}$$

Based on the model specifications above, Table 2 summarizes the expected direction of influence of this paper's core explanatory variables and their interaction terms.

Table 2: Expected (Sign) Behavior for Core Variables

Core Variables	Expected Sign	Core Variables	Expected Sign
lnTextDigital	+	ROD2	+
L1.lnTextDigital	+/-	L1.ROD2	+
lnTextDigital × EU	+	ROD2 × EU	+
lnTextDigital × During	+/-	ROD2 × During	+
lnTextDigital × Post	+/-	ROD2 × Post	+
lnTextDigital × Stringency_during	+/-	ROD2 × Stringency_during	+

Source: Author's elaboration

3.3.5 Robustness Checks and Supplementary Analyses

To test the robustness of the main conclusions and systematically evaluate the validity of the ROD2 proxy variable, this paper conducts the set of checks summarized in Table 3. The complete results are reported in Appendix Tables 8-20.

Table 3: Robustness Checks and Supplementary Analysis Framework

Test item	Objective	Expected result	Interpretation
Relationship between lnTextDigital and ROD2	Verify whether ROD2 is empirically associated with textual digital attention	Positive and significant	Provides suggestive support for the operational conversion interpretation of ROD2, but does not imply that ROD2 is a digital specific measure
Industry subgroup test by digital intensity	Test whether ROD2 has some digital related relevance	Significant in high digital intensity industries	If not confirmed, ROD2 is more likely to reflect general operating efficiency
lnTextDigital × ROD2 interaction	Test complementarity between textual digital attention and operational conversion capability	Positive and significant	If insignificant, complementarity is not consistently supported
Profit Margin control	Mitigate DuPont related overlap between ROD2 and ROA or ROE	ROD2 results remain stable	Reduces, but does not eliminate, the accounting overlap concern
Cash Flow Ratio as an alternative performance indicator	Reduce reliance on profit based performance measures	ROD2 remains significant	Supports the relevance of ROD2 beyond ROA and ROE
Std_lnTextDigital replacing lnTextDigital	Control for differences in annual report length	Results remain stable	Indicates that textual indicator results are not mainly driven by report length
lnROD replacing ROD2	Test the robustness of the operational conversion indicator	Results remain stable	Supports the robustness of the operational conversion dimension
CountryStringency_during replacing Stringency_during	Test the matching scheme for the policy stringency variable	Results remain stable	Indicates that policy stringency findings do not depend on the specific matching level
1% winsorization	Reduce the influence of extreme values	Results remain similar	Supports robustness to extreme observations
Industry × year fixed effects	Control for time varying industry level shocks	Results remain stable	Indicates that the findings are not mainly driven by common industry level shocks
Future period digital transformation indicators	Check potential reverse causality or temporal persistence	Insignificant	If significant, the direction of the results should be interpreted with caution

Source: Author’s elaboration

4. Results

4.1 Descriptive Analysis

Table 4 reports descriptive statistics for the full sample as well as the EU and China subsamples. The final sample comprises 25,120 firm-year observations, covering listed companies in mainland China and the EU from 2015 to 2024. For the full sample, the mean ROA is 3.2449, the mean ROE is 4.5535, and the mean Tobin's Q is 1.8699; the core explanatory variables lnTextDigital and ROD2 have means of 3.1862 and 0.6093, respectively. Some variables exhibit considerable dispersion. For example, the maximum value of Price_to_Book_Ratio reaches 3254.8609, while the minimum value of ROE is -5065.6653, indicating the presence of fairly extreme observations in the sample.

In terms of regional differences, the mean ROA is similar between the EU and China samples, but the mean ROE for the EU sample is 8.3972, markedly higher than the 3.3814 observed for the China sample. With respect to the digitalization variables, the mean lnTextDigital for the China sample is 3.2697, higher than the EU sample's 2.9125; however, the mean ROD2 for the China sample is 0.5871, lower than the EU sample's 0.6819, indicating some difference between the two regions in textual digital attention and asset output efficiency. For the pandemic variables, COVID_Shock and Stringency_Index cover 7,536 observations from 2020 to 2022. The EU sample shows a higher average COVID_Shock, while the China sample shows a higher average Stringency_Index. This pattern suggests that infection based and policy based pandemic measures capture different dimensions of the COVID-19 context.

Table 4: Descriptive Statistics: EU, China and Full Sample

Variable	Region	Obs.	Mean	Std. dev.	Min	Max
ROA	Total	25,120	3.2449	8.5530	-297.2635	181.1679
	EU	5,870	3.2311	12.3437	-297.2635	181.1679
	China	19,250	3.2491	7.0003	-164.7913	78.5865
ROE	Total	25,120	4.5535	53.8574	-5065.6653	1746.2600
	EU	5,870	8.3972	46.7317	-966.8600	1746.2600
	China	19,250	3.3814	55.7982	-5065.6653	233.1103
Tobin's Q	Total	25,120	1.8699	1.4671	0.2847	56.6643
	EU	5,870	1.4920	1.0755	0.2847	15.2386
	China	19,250	1.9852	1.5490	0.6112	56.6643
lnTextDigital	Total	25,120	3.1862	1.1995	0.0000	6.8865
	EU	5,870	2.9125	1.3060	0.0000	6.4052
	China	19,250	3.2697	1.1523	0.0000	6.8865

ROD2	Total	25,120	0.6093	0.5101	-0.0502	11.4110
	EU	5,870	0.6819	0.5341	0.0000	6.1933
	China	19,250	0.5871	0.5005	-0.0502	11.4110
COVID_Shock	Total	7536	0.0330	0.0860	0.0000	0.4294
	EU	1761	0.1403	0.1288	0.0071	0.4294
	China	5775	0.0002	0.0005	0.0000	0.0025
Stringency_Index	Total	7536	52.2079	11.6382	12.4903	74.3655
	EU	1761	42.3894	16.9536	12.4903	72.2702
	China	5775	55.2019	7.1260	37.9722	74.3655
Size	Total	25,120	7.9821	1.5498	0.0000	13.7839
	EU	5,870	7.9265	2.1599	0.0000	13.0082
	China	19,250	7.9990	1.3080	2.9957	13.7839
Leverage	Total	25,120	0.4780	0.2160	0.0042	3.0131
	EU	5,870	0.5965	0.2181	0.0042	3.0131
	China	19,250	0.4418	0.2020	0.0131	0.9974
Cash_Flow_Ratio	Total	25,120	0.0548	0.3074	-36.6202	29.3992
	EU	5,870	0.0630	0.6239	-36.6202	29.3992
	China	19,250	0.0524	0.0678	-0.9129	0.8385
Fixed_Assets_Ratio	Total	25,120	0.2191	0.1836	0.0000	0.9901
	EU	5,870	0.2573	0.2391	0.0000	0.9901
	China	19,250	0.2074	0.1611	0.0002	0.9542
Price_to_Book_Ratio	Total	25,120	3.2192	23.1314	0.0266	3254.8609
	EU	5,870	2.3498	3.9925	0.0266	185.9732
	China	19,250	3.4843	26.3262	0.1077	3254.8609

Notes: N refers to firm-year observations. COVID_Shock and Stringency_Index are observed only for 2020-2022.
Source: Author's elaboration

Model specification tests further support the empirical design. The variance inflation factor (VIF) values are all below 2.1, well below the commonly used threshold of 10 (Meng *et al.*, 2017). The core variables pass the panel unit root tests in levels, and the F tests and Hausman tests support the use of firm and year fixed effects. The full results are reported in Appendix Tables 3-7.

4.2 Baseline Results and Lagged Effects

Table 5 reports the baseline regression results using firm and year fixed effects, with ROA, ROE, and Tobin's Q as dependent variables. Panel A and Panel B correspond to lnTextDigital and ROD2, respectively.

Panel A shows that lnTextDigital is significantly positive for both ROA and ROE, with coefficients of 0.3400 and 2.0242, respectively. This indicates that textual digital attention is associated with better contemporaneous accounting performance. However, its effect on Tobin's Q is not significant, suggesting that textual digital attention does not show a stable positive relationship with contemporaneous market valuation.

Panel B shows that ROD2 is significantly positive for ROA and Tobin's Q, with coefficients of 2.3458 and 0.3679, respectively, while its effect on ROE is positive but not significant. This suggests that higher asset output efficiency is associated with better ROA and stronger market valuation performance.

Table 5: Baseline Regression Results

Variables	ROA	ROE	Tobin's Q	ROA	ROE	Tobin's Q
	Panel A: lnTextDigital			Panel B: ROD2		
lnTextDigital	0.3400***	2.0242**	-0.0203			
	(0.0954)	(1.0005)	(0.0179)			
ROD2				2.3458*	0.1351	0.3679***
				(1.2566)	(5.1381)	(0.0546)
Size	0.8981	4.0083**	-0.1940***	0.8955	4.5508***	-0.2143***
	(0.6366)	(1.6907)	(0.0378)	(0.6339)	(1.6004)	(0.0378)
Leverage	-21.0642***	-58.0572***	0.0406	-21.0311***	-58.2939***	0.0542
	(1.3913)	(10.7711)	(0.1535)	(1.4015)	(10.8281)	(0.1527)
Cash Flow Ratio	1.1799	1.3362	-0.0012	1.1419	1.3269	-0.0069
	(0.9051)	(1.3316)	(0.0239)	(0.8798)	(1.3311)	(0.0220)
Fixed Assets Ratio	-11.1362***	-7.9523	0.2899	-11.4121***	-9.2556	0.2938
	(1.0527)	(18.3367)	(0.2888)	(1.0527)	(17.7832)	(0.2871)
Price-to-Book	0.0012	-1.4650***		0.0010	-1.4653***	
	(0.0018)	(0.1272)		(0.0018)	(0.1273)	
Constant	7.4316	0.2445	3.4002***	7.1537	2.6810	3.2664***
	(4.7740)	(11.2310)	(0.2805)	(4.7735)	(11.7410)	(0.2816)
Observations	25,120	25,120	25,120	25,120	25,120	25,120
Adjusted R^2	0.4253	0.4562	0.6041	0.4282	0.4558	0.6067

Notes: Robust standard errors clustered at the firm level are reported in parentheses. Firm and year fixed effects are included. Price-to-Book Ratio is excluded from the Tobin's Q models. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.
Source: Author's elaboration

Table 6 further reports the results using digitalization variables lagged by one period. Panel A shows that lagged lnTextDigital is significantly negative for ROA and Tobin's Q, with coefficients of -0.2740 and -0.0283, respectively, while its effect on ROE is not statistically significant. This suggests that textual digital attention does not consistently translate into improved subsequent performance, and may even be associated with weaker subsequent ROA and market valuation. By contrast, Panel B shows that lagged ROD2 is significantly positive for ROA, ROE, and Tobin's Q, indicating that the asset output efficiency and operational conversion capacity reflected by ROD2 have a more stable positive relationship with subsequent firm performance.

Overall, the baseline and lagged results support H1. Textual digital attention is positively associated with contemporaneous ROA and ROE, but this relationship does not persist in the lagged models. ROD2, by contrast, shows a stronger and more stable positive association with firm performance over time.

Table 6: Lagged Regression Results

Variables	ROA	ROE	Tobin's Q	ROA	ROE	Tobin's Q
	Panel A: L1.lnTextDigital			Panel B: L1.ROD2		
L1.lnTextDigital	-0.2740**	0.2173	-0.0283**			
	(0.1179)	(0.9775)	(0.0142)			
L1.ROD2				2.4910***	5.7113***	0.1969***
				(0.4296)	(1.4317)	(0.0341)
Size	1.0888	5.5381***	-0.1744***	0.9019	5.3240***	-0.1909***
	(0.7330)	(2.0264)	(0.0328)	(0.7136)	(1.8940)	(0.0330)
Leverage	-22.0183***	-63.3337***	0.1700	-21.7989***	-62.8595***	0.1877
	(1.4686)	(11.6856)	(0.1334)	(1.4490)	(11.6781)	(0.1333)
Cash Flow Ratio	1.7195	2.1441	0.0289	1.6781	2.0430	0.0257
	(1.7277)	(2.5906)	(0.0305)	(1.7031)	(2.5315)	(0.0288)
Fixed Assets Ratio	-12.0634***	-7.2761	0.1763	-11.7465***	-6.9610	0.2046
	(1.1780)	(22.4980)	(0.2039)	(1.1763)	(22.1228)	(0.2044)
Price-to-Book	0.0010	-1.4712***		0.0010	-1.4712***	
	(0.0020)	(0.1238)		(0.0019)	(0.1241)	
Constant	8.4471	-4.3123	3.1486***	7.3882	-5.7021	3.0567***
	(5.4671)	(13.2692)	(0.2461)	(5.4253)	(13.5480)	(0.2475)
Observations	22,608	22,608	22,608	22,608	22,608	22,608
Adjusted R^2	0.4293	0.4627	0.6440	0.4325	0.4631	0.6448

Notes: Robust standard errors clustered at the firm level are reported in parentheses. Firm and year fixed effects are included. Price-to-Book Ratio is excluded from the Tobin's Q models. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.
Source: Author's elaboration

4.3 Institutional Heterogeneity Analysis

Table 7 reports the results from models incorporating interaction terms between the digitalization indicators and the EU dummy variable, which are used to examine whether institutional differences moderate the relationship between digital transformation related indicators and firm performance.

Panel A shows that lnTextDigital × EU is significantly positive for ROA, ROE, and Tobin's Q, with coefficients of 0.8291, 3.5187, and 0.2472, respectively. This indicates that the marginal relationship between textual digital attention and firm performance is stronger in the selected EU sample than in the China sample, covering both accounting performance and market valuation.

Panel B shows that ROD2 × EU is significantly positive for ROA and ROE, with coefficients of 4.4052 and 22.9430, respectively, but is not significant for Tobin's Q. This suggests that the institutional difference associated with ROD2 is mainly reflected in accounting performance rather than market valuation.

Overall, the results indicate that the marginal associations between digitalization indicators and firm performance vary across the China and selected EU samples. Within this EU subsample, textual digital attention shows a stronger marginal relationship with firm performance, while the asset output efficiency captured by ROD2 is more strongly associated with accounting performance. Taken together, these findings support H2.

Table 7: Institutional Differences in Digitalization Effects

Variables	ROA	ROE	Tobin's Q	ROA	ROE	Tobin's Q
	Panel A: lnTextDigital × EU			Panel B: ROD2 × EU		
lnTextDigital	0.1254	1.1132	-0.0843***			
	(0.1133)	(1.2657)	(0.0220)			
lnTextDigital × EU	0.8291***	3.5187**	0.2472***			
	(0.2003)	(1.4338)	(0.0287)			
ROD2				1.5806	-3.8500	0.3477***
				(1.4293)	(5.6892)	(0.0590)
ROD2 × EU				4.4052**	22.9430***	0.1161
				(2.0818)	(8.0813)	(0.1376)
Size	0.9306	4.1462**	-0.1842***	0.8918	4.5315***	-0.2144***
	(0.6418)	(1.7295)	(0.0375)	(0.6333)	(1.6033)	(0.0378)
Leverage	-20.9852***	-57.7222***	0.0634	-21.0239***	-58.2567***	0.0544
	(1.3833)	(10.8050)	(0.1534)	(1.4038)	(10.7193)	(0.1526)
Cash Flow Ratio	1.1833	1.3507	-0.0001	1.1357	1.2948	-0.0070
	(0.9036)	(1.3264)	(0.0236)	(0.8732)	(1.2969)	(0.0218)
Fixed Assets Ratio	-11.3158***	-8.7146	0.2362	-11.0162***	-7.1934	0.3043
	(1.0506)	(18.5506)	(0.2897)	(1.0620)	(17.8599)	(0.2882)
Price-to-Book	0.0011	-1.4656***		0.0010	-1.4654***	
	(0.0019)	(0.1270)		(0.0018)	(0.1271)	
Constant	7.2937	-0.3407	3.3587***	6.8579	1.1403	3.2586***
	(4.8029)	(11.3853)	(0.2765)	(4.7629)	(11.8193)	(0.2811)

Observations	25,120	25,120	25,120	25,120	25,120	25,120
Adjusted R^2	0.4260	0.4565	0.6061	0.4297	0.4569	0.6068

Notes: Robust standard errors clustered at the firm level are reported in parentheses. Firm and year fixed effects are included. Price-to-Book Ratio is excluded from the Tobin's Q models. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.
Source: Author's elaboration

4.4 COVID-19 and Policy Stringency

This section examines whether the COVID-19 context altered the relationship between digital transformation and firm performance from two perspectives: pandemic stage heterogeneity and policy constraint intensity.

4.4.1 Pandemic-Stage Heterogeneity

Table 8 uses the period before the pandemic, 2015-2019, as the reference group to examine whether the relationship between digitalization variables and performance changed during and after the pandemic.

Panel A shows that the main effects of lnTextDigital on ROA and ROE before the pandemic are broadly consistent with the baseline regression results. However, the interaction terms between lnTextDigital and the pandemic stage dummy variables are not statistically significant across the three performance indicators. This suggests that the pandemic context did not systematically strengthen the relationship between textual digital attention and firm performance.

Panel B shows a different pattern. Before the pandemic, the main effect of ROD2 is significant only for Tobin's Q, which is consistent with the baseline regression results. However, the coefficients of ROD2 × During COVID on ROA and ROE are 1.4156 and 8.3496, respectively, and both are significant at the 1% level. ROD2 × Post COVID is also significantly positive for ROA and ROE, with coefficients of 1.2340 and 9.1183. The similar magnitude of the interaction coefficients during and after the pandemic indicates that the asset output efficiency reflected by ROD2 has a strengthened relationship with accounting performance after the onset of the pandemic, rather than only during the most severe stage of the crisis. Overall, in a crisis context, firms' ability to convert resources into operating output has stronger explanatory power for performance than strategic statements about digitalization.

Table 8: Pandemic Stage Heterogeneity in Digitalization Effects

Variables	ROA	ROE	Tobin's Q	ROA	ROE	Tobin's Q
	Panel A: lnTextDigital × COVID stages			Panel B: ROD2 × COVID stages		
lnTextDigital	0.3743***	2.0569*	-0.0175			
	(0.1004)	(1.0691)	(0.0207)			
ROD2				1.7420	-3.6734	0.3450***
				(1.3523)	(5.2598)	(0.0570)
lnTextDigital × During-COVID	-0.0813	-0.7286	-0.0257			
	(0.1558)	(0.6996)	(0.0188)			
lnTextDigital × Post-COVID	-0.1158	1.0099	0.0232			
	(0.1297)	(0.6304)	(0.0222)			
ROD2 × During-COVID				1.4156***	8.3496***	0.0500
				(0.4494)	(2.8475)	(0.0347)
ROD2 × Post-COVID				1.2340***	9.1183***	0.0548
				(0.4157)	(2.5736)	(0.0348)
Size	0.9069	3.9726**	-0.1946***	0.8713	4.3714***	-0.2154***
	(0.6367)	(1.6768)	(0.0379)	(0.6352)	(1.6154)	(0.0377)
Leverage	-21.0428***	-58.0146***	0.0431	-20.9776***	-57.8774***	0.0567
	(1.4041)	(10.8266)	(0.1537)	(1.3893)	(10.8032)	(0.1527)
Cash Flow Ratio	1.1815	1.3319	-0.0012	1.1288	1.2389	-0.0074
	(0.9061)	(1.3306)	(0.0239)	(0.8780)	(1.3218)	(0.0218)
Fixed Assets Ratio	-11.1407***	-7.9373	0.2901	-11.4139***	-9.2721	0.2937
	(1.0508)	(18.3289)	(0.2887)	(1.0504)	(17.7223)	(0.2870)
Price-to-Book Ratio	0.0012	-1.4649***		0.0010	-1.4654***	
	(0.0018)	(0.1272)		(0.0018)	(0.1270)	
Constant	7.4094	0.4243	3.4043***	7.2845	3.6288	3.2720***
	(4.7745)	(11.2358)	(0.2811)	(4.7949)	(11.8621)	(0.2816)
Observations	25,120	25,120	25,120	25,120	25,120	25,120
Adj. R-squared	0.4253	0.4563	0.6043	0.4298	0.4575	0.6068

Notes: Pandemic-stage main effects are absorbed by year fixed effects because they are defined by calendar years. Robust standard errors clustered at the firm level are reported in parentheses. Firm and year fixed effects are included. Price-to-Book Ratio is excluded from the Tobin's Q models. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.
Source: Author's elaboration

4.4.2 Policy Stringency and the Performance Effects of Digitalization

Table 9 further examines the moderating effect of policy constraint intensity during the pandemic. In Panel A, lnTextDigital × Stringency_during is not significant for either ROA or ROE, while it shows only a weakly significant negative effect on Tobin's Q. This suggests that, under stricter policy constraints during the pandemic, textual digital strategy disclosure did not provide additional support for market valuation.

Panel B shows that Stringency_during is significantly negative for ROA, ROE, and Tobin's Q, indicating that stricter containment policies placed clear pressure on firm performance. At the same time, ROD2 × Stringency_during is significantly positive for ROA and ROE, with coefficients of 0.0187 and 0.1095, respectively. This suggests that firms with higher asset output efficiency show a certain buffering advantage in accounting performance under stronger policy constraints. However, this buffering effect does not extend to market valuation, indicating that the role of ROD2 in the pandemic context is mainly reflected in accounting performance.

Table 9: Policy Stringency and the Performance Effects of Digitalization

Variables	ROA	ROE	Tobin's Q	ROA	ROE	Tobin's Q
	Panel A: lnTextDigital × policy			Panel B: ROD2 × policy stringency		
lnTextDigital	0.3703***	2.1527**	-0.0125			
	(0.0951)	(1.0341)	(0.0191)			
ROD2				2.0354	-1.6869	0.3591***
				(1.3149)	(5.2601)	(0.0557)
Stringency_during	-0.0101	-0.1087	-0.0011	-0.0260**	-0.1854**	-0.0034***
	(0.0113)	(0.0924)	(0.0013)	(0.0109)	(0.0919)	(0.0007)
lnTextDigital × Stringency_during	-0.0020	-0.0066	-0.0006*			
	(0.0021)	(0.0075)	(0.0003)			
ROD2 × Stringency_during				0.0187***	0.1095***	0.0005
				(0.0066)	(0.0388)	(0.0005)
Size	0.9040	4.0401**	-0.1926***	0.8995	4.5835***	-0.2134***
	(0.6356)	(1.6841)	(0.0377)	(0.6337)	(1.5928)	(0.0378)
Leverage	-21.0413***	-57.9628***	0.0465	-21.0384***	-58.3299***	0.0545
	(1.3963)	(10.7809)	(0.1534)	(1.4008)	(10.8445)	(0.1526)
Cash Flow Ratio	1.1810	1.3421	-0.0009	1.1400	1.3167	-0.0068
	(0.9073)	(1.3452)	(0.0239)	(0.8811)	(1.3418)	(0.0219)
Fixed Assets Ratio	-11.1480***	-8.0289	0.2875	-11.4115***	-9.2769	0.2918
	(1.0500)	(18.3218)	(0.2888)	(1.0467)	(17.7522)	(0.2868)
Price-to-Book Ratio	0.0012	-1.4652***		0.0010	-1.4656***	
	(0.0018)	(0.1272)		(0.0018)	(0.1272)	
Constant	7.5461	1.6101	3.4096***	7.5462	5.4279	3.3136***
	(4.7800)	(10.9127)	(0.2813)	(4.7831)	(11.5528)	(0.2816)
Observations	25,120	25,120	25,120	25,120	25,120	25,120
Adj. R-squared	0.4255	0.4564	0.6044	0.4290	0.4567	0.6069

Notes: Robust standard errors clustered at the firm level are reported in parentheses. Firm and year fixed effects are included. Price-to-Book Ratio is excluded from the Tobin's Q models. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.
Source: Author's elaboration

4.4.3 Supplementary Tests of Pandemic-Context Variables

As a supplementary test, this paper uses COVIDShock_during, constructed from confirmed case counts, to examine the role of pandemic transmission intensity. The interaction terms between COVIDShock_during and the two digitalization indicators are not significant across the ROA, ROE, and Tobin's Q models, and the results are reported in Appendix Table 22. The mean difference tests reported in Appendix Table 21 further show that COVID_Shock and Stringency_Index differ significantly in their regional patterns across the China and EU samples. Therefore, this paper mainly relies on the policy stringency indicator to interpret the buffering effect observed in the pandemic context.

Collectively, the COVID-19 analysis shows that textual digital attention does not provide a stable additional performance advantage during the pandemic or under stronger policy constraints. By contrast, the asset output efficiency and operational conversion capability reflected by ROD2 are more strongly associated with accounting performance in the pandemic context. Therefore, H3 is partially supported, with the supporting evidence concentrated in accounting performance.

4.5 Robustness Checks

This section summarizes a series of robustness checks and supplementary analyses.

First, this study conducts supplementary tests to clarify the interpretation of ROD2 as an asset output efficiency indicator and an indirect proxy for operational conversion capacity. The association between lnTextDigital and ROD2 is reported in Appendix Table 8. The industry digital intensity test provides suggestive evidence that the performance relevance of ROD2 is stronger in more digitally intensive industries. This supports a cautious interpretation of ROD2 as an indirect proxy for operational conversion capacity related to digital transformation. Additional tests, including Profit Margin, Cash Flow Ratio, industry digital intensity subgroups, and the lnTextDigital × ROD2 interaction, are reported in Appendix Tables 9-13. These tests are used to further clarify the interpretive boundaries of ROD2. The insignificant interaction term further suggests that textual digital attention and asset output efficiency are better interpreted as separate dimensions rather than as mutually reinforcing mechanisms.

Second, the main conclusions remain broadly stable when alternative variables are used. Replacing lnTextDigital with Std_lnTextDigital shows that the textual digitalization results are not mainly driven by differences in annual report length. After further excluding broad terms such as integration, information system, and networked, the H1 results remain broadly consistent, suggesting that the accounting performance results of the text based digitalization indicator are not entirely driven by generic informatization terms. Replacing ROD2 with lnROD produces results that are broadly consistent for accounting performance, although differences appear in the Tobin's Q model. This suggests that asset output efficiency and labor output efficiency may not follow the same path in relation to market valuation. In addition, using CountryStringency_during instead of Stringency_during produces similar results, indicating that the policy stringency findings do not depend on the specific regional matching scheme. These alternative measure and policy stringency checks are reported in Appendix Tables 14-16 and 20.

Third, the results remain stable after alternative sample processing and model specifications. After applying 1% winsorization, the main findings for lnTextDigital and ROD2 remain broadly consistent. After adding industry × year fixed effects, the key conclusions also remain stable, suggesting that the results are not mainly driven by extreme observations or time varying industry level shocks. These checks are reported in Appendix Tables 17-18.

Finally, the temporal placebo test using future period digitalization variables suggests that some associations may reflect temporal persistence or a potential bidirectional relationship between digitalization indicators and firm performance. Therefore, the results should be interpreted mainly as empirical associations and marginal differences across digitalization dimensions, rather than as strict causal effects (Appendix Table 19).

Overall, the robustness checks support the main conclusion that different dimensions of digital transformation are associated with firm performance in different ways. The core findings are not driven by a specific indicator construction method, extreme observations, or structural overlap among accounting indicators.

5. Discussion

5.1 Theoretical Interpretation

The empirical results reveal three main aspects of the relationship between digital transformation and firm performance: differences between pathways, institutional boundary conditions, and the crisis context. They also provide several supplementary findings relative to existing theoretical frameworks.

Regarding pathway differences and performance dynamics, lnTextDigital and ROD2 show substantively different relationships with firm performance. The contemporaneous coefficient of lnTextDigital is significantly positive for both ROA and ROE, whereas its coefficient lagged by one period turns significantly negative for ROA and Tobin's Q. This pattern suggests that digital attention disclosed in annual reports reflects firms' strategic agenda and disclosure intent more than a realized improvement in operational efficiency. From the perspective of disclosure theory, management statements about digitalization may help reduce information asymmetry in the current period. However, the credibility of such signals depends on whether the underlying capabilities can be verified, which makes it difficult for these signals to translate consistently into performance improvement in later periods (Healy & Palepu, 2001; Verrecchia, 2001). Brynjolfsson's (1993) argument on the information technology productivity paradox further suggests that digital technology investment does not immediately lead to performance improvement. Firms usually need complementary investments, including organizational adjustment, employee training, and process redesign, before technological resources can be converted into observable efficiency gains (Brynjolfsson & Hitt, 2000). So, a higher level of digitalization disclosure in the previous period may indicate that the firm is still in an investment heavy stage, where costs appear before benefits, thereby weakening accounting performance in the subsequent period.

By contrast, ROD2 exhibits a more stable positive performance effect in both the contemporaneous and one period lagged models, and remains significant for ROA, ROE, and Tobin's Q in the lagged models. From a DCV perspective, the asset output efficiency reflected by ROD2 can be understood as an operational manifestation of firms' resource reconfiguration and allocation optimization (Teece *et al.*, 1997; Helfat & Winter, 2011).

This is consistent with Yavuz *et al.*'s (2025) finding that operational efficiency indicators have more stable predictive power for performance. It is worth noting that the contemporaneous effect of ROD2 on ROE is not significant, whereas it becomes significant when lagged by one period. This may be because ROE is affected not only by operating efficiency, but also by capital structure and financial leverage (Bunea *et al.*, 2019). Therefore, the transmission from operational efficiency improvement to shareholder returns may take longer. In addition, given that ROD2 has a certain accounting linkage with ROA and ROE, its economic meaning should not be interpreted based on accounting performance results alone. The results show that ROD2 also maintains a significant positive relationship in the Tobin's Q market valuation model. Moreover, in supplementary checks that further control for Profit Margin and use Cash Flow Ratio as an alternative performance measure, the main conclusions remain broadly stable. This indicates that the explanatory power of ROD2 is not limited to accounting performance indicators, providing a more robust basis for interpreting it as a proxy for asset output efficiency and operational conversion capacity.

Taken together, the relationship between text based digital attention and firm performance is not stable over time, whereas the asset output efficiency reflected by ROD2 shows a more stable positive relationship with subsequent performance. This indicates that, compared with strategic level digital expression, operational level efficiency improvement and resource conversion capacity better explain intertemporal differences in firm performance.

The insignificant interaction between lnTextDigital and ROD2 indicates only a limited complementary relationship between textual digital attention and asset output efficiency in the sample. This helps clarify that the two indicators are related but relatively independent, with lnTextDigital capturing strategic attention and ROD2 serving as an indirect operational conversion indicator based on operating results.

With respect to institutional boundary conditions, the interaction analysis shows that the selected EU sample displays stronger marginal associations between digitalization indicators and firm performance, providing support for H2. lnTextDigital × EU is significantly positive for ROA, ROE, and Tobin's Q, whereas the significance of ROD2 × EU is concentrated in the two accounting performance indicators, ROA and ROE, and is not significant for Tobin's Q. This channel selectivity is consistent with differences between China and the EU in information ecosystems and market rules. The EU's unified digital market framework and data governance rules, exemplified by the GDPR, may provide a more predictable disclosure and market environment in which indicators related to digitalization are more strongly associated with performance (Cai & Zhang, 2022; Liang & Wu, 2020). In such an environment, market participants may find it easier to interpret firms' digital strategy statements as verifiable capability commitments, while efficiency gains may be more closely associated with accounting performance through scale expansion under unified market rules. However, after applying the two screening criteria of complete English annual report disclosure and complete panel data, only part of the original EU firm sample is retained. Systematic differences in industry structure, firm size, internationalization, and disclosure quality may also amplify the marginal relationship between digital transformation indicators and firm performance in the EU sample. Under the current research design, institutional environment effects and sample composition effects cannot be fully separated.

In the China context, the significantly negative effect of lnTextDigital on Tobin's Q and the insignificant results of ROD2 for some performance indicators suggest that the conversion of digital strategic attention and operational efficiency into performance is relatively less stable. China's digital market rules remain in a stage of dynamic evolution, and the regulatory framework relies more on flexible policy guidance than on stable hard law rules (Zheng & Snyder, 2023). Digital governance also places greater emphasis on security orientation and state regulation than on unified market competition rules (Wang & Han, 2026). Given this lower degree of rule stability, market participants may find it more difficult to independently verify firms' digitalization disclosures. As a result, they may interpret extensive digitalization related language in annual reports as strategic statements that have not yet been realized, which may help explain its negative association with market valuation. This observed difference in the selected EU sample is not evenly distributed across channels. It suggests that information ecosystems may matter more for the credibility of textual disclosure, while market rules may matter more for the link between operational efficiency and accounting performance. The two institutional mechanisms therefore operate through different channels.

Regarding crisis context effects, the pandemic analysis provides partial support for H3 and further reveals a divergence between the two types of digitalization indicators under crisis conditions. The interaction terms between lnTextDigital and the pandemic period and post pandemic stages are not statistically significant across the three performance indicators. This indicates that the degree of digital attention disclosed in annual reports did not provide a stronger performance advantage during the pandemic context. This is consistent with Chen *et al.*'s (2024) view that the contribution of digital transformation to corporate resilience depends on whether technology is genuinely embedded in operating processes. By contrast, the interaction terms between ROD2 and the pandemic period and post pandemic stages show significant positive effects on both ROA and ROE, indicating that firms with higher asset output efficiency were better able to convert operational advantages into relative performance gains under external shocks.

The baseline effect of policy stringency is negative for ROA, ROE, and Tobin's Q, reflecting the direct pressure that containment policies placed on firm performance. The significant positive results for ROD2 × Stringency_during on ROA and ROE further indicate that, under stronger policy constraints, firms with higher asset output efficiency achieved relatively stronger accounting performance (Bai *et al.*, 2021; Härting *et al.*, 2022; Chen *et al.*, 2024). From a DCV perspective, the COVID-19 shock, as a highly uncertain exogenous disturbance, directly tests firms' capacity to seize opportunities and reconfigure resources (Teece *et al.*, 1997). The insignificant results for the confirmed case interaction term likewise indicate that the intensity of policy control, rather than the scale of infection, better captures the operating constraints that firms actually faced. This is consistent with the fundamental differences in pandemic containment approaches between China and the EU.

Overall, the theoretical contribution of this paper lies in revealing the dimensional differentiation, temporal dynamics, and contextual dependence of the performance effects of digital transformation. The economic meanings captured by the strategic expression path and the operational conversion path differ substantively, and the digitalization paradox more likely reflects stage specific differences in measurement rather than an inherent limitation of transformation itself. The moderating roles of the institutional environment and the crisis context further indicate that the findings are consistent with the view that the performance relevance of digitalization indicators is shaped not only by firm-level characteristics, but also by external regulatory frameworks and contextual pressures. This perspective provides a more refined analytical foundation for future research on indicator construction, contextual modeling, and causal identification.

5.2 Implications

For corporate managers, the results indicate that digital transformation should not remain in strategic statements and annual report disclosure. ROD2 shows stable predictive power in contemporaneous and lagged models, suggesting that asset output efficiency, as an indirect indicator of operational conversion, is more closely associated with performance than textual digital attention. Firms should therefore ensure that digital technologies are genuinely embedded in production, sales, supply chain management, customer management, and internal decision making, rather than simply increasing the frequency of digitalization related language in annual reports. The pandemic context findings further show that operational digital capability provides more direct support for corporate resilience. Firms should proactively build digital operational infrastructure, including supply chain collaboration platforms, remote operation systems, and real time business data monitoring capabilities, to improve resource allocation flexibility when facing external shocks. When evaluating transformation outcomes, firms should consider accounting performance, market valuation, and operational efficiency together, and establish a dynamic tracking mechanism across multiple financial periods to avoid misjudging transformation progress based on short term performance fluctuations.

For policymakers, the results suggest that regulators should guide firms to disclose more specific and verifiable digitalization information, including the scale of digital investment, infrastructure development, digital talent deployment, and the scope of process coverage. This would help reduce market uncertainty about the credibility of firms' digital strategies and shift policy support from strategic digitalization rhetoric toward substantive digital capability building. Digital resilience should also be incorporated into the long term economic governance framework. Policy instruments could include digital operational capability building within public procurement, financial support, and industrial subsidies, so as to strengthen the overall resilience of the economic system against future external shocks, such as public health events, supply chain disruptions, and geopolitical risks.

6. Conclusions, Limitations and Future Research

6.1 Conclusions

This study uses listed companies in China and the EU from 2015 to 2024 as the research sample. It constructs the text based digitalization indicator lnTextDigital and the asset output efficiency indicator ROD2 to examine the relationship between digital transformation and firm performance from the perspectives of strategic attention and operational conversion. It also considers three performance indicators, namely ROA, ROE, and Tobin's Q, while incorporating China and EU institutional differences and the COVID-19 shock into the analysis.

The empirical analysis yields three main conclusions. First, the relationship between digital transformation and firm performance shows clear dimensional differences and temporal dynamics. lnTextDigital is mainly associated with contemporaneous improvement in accounting performance, while its lagged relationship is unstable and turns negative for some performance indicators, which is consistent with the information technology productivity paradox. By contrast, ROD2 shows a more stable positive relationship in both contemporaneous and lagged models, suggesting that asset output efficiency and operational conversion capacity have stronger explanatory power for performance over time. Second, the institutional environment is an important boundary condition. The marginal relationships of both digitalization indicators are generally stronger in the selected EU sample, but the observed difference varies between the textual disclosure and operational efficiency channels. This suggests that information ecosystems and market rules may shape the observed associations between digitalization indicators and firm performance through different channels. Third, the pandemic context further highlights the importance of operational efficiency. The interactions of ROD2 with the pandemic period and policy stringency are positively associated with accounting performance, whereas the corresponding interactions for lnTextDigital do not show stable effects. Under external shocks, the asset output efficiency and operational conversion capacity reflected by ROD2 may therefore help explain differences in firm resilience.

This study provides supplementary empirical evidence on how, when, and under what conditions digital transformation is related to firm performance. The findings suggest that, compared with strategic statements about digitalization, operational conversion capacity reflected in asset output efficiency and resource allocation is more consistently associated with persistent differences in firm performance.

6.2 Limitations and Future Research

This study has several limitations. The EU sample includes only listed firms with complete English annual reports, which may bias its geographic distribution and industry structure. The conclusions are therefore more applicable to firms with greater internationalization and more complete disclosure. Cross language keyword matching may also introduce bias in Chinese-English semantic equivalence. The three stage control process helps reduce this issue but cannot fully remove it. In addition, ROD2 has a certain accounting linkage with ROA and ROE. Although this study addresses the issue through alternative performance measures and additional controls, this linkage remains an important boundary condition. The use of a balanced panel may also bias the sample toward larger and more stable listed firms, creating some survivorship bias. Finally, endogeneity remains a major challenge. Two way fixed effects and lagged variables can only partly reduce the influence of omitted variables and reverse causality.

Future research could be extended in several directions. At the sample level, multilingual text analysis or an unbalanced panel could be used to improve the representativeness of the EU sample and examine survivorship bias. At the indicator level, more direct measures of digitalization, such as IT investment, digital patents, or digital talent, could be included. Researchers could also construct efficiency variables that are orthogonal to asset size to better exclude the influence of accounting structure. At the causal identification level, instrumental variables, difference in differences, or quasi natural experiments could strengthen causal inference, for example by using digitalization policies or changes in data governance rules as external shocks. At the industry level, future studies could examine the textual disclosure path and operational conversion path under more detailed industry classifications, thereby identifying more precisely how digital transformation creates value across different industries.

References

Al-Matari, E. M., Al-Swidi, A. K., & Fadzil, F. H. B. (2014). The measurements of firm performance's dimensions. *Asian Journal of Finance & Accounting*, *6*(1), 24-49. https://doi.org/10.5296/ajfa.v6i1.4761

Axenbeck, J., & Breithaupt, P. (2022). *Measuring the digitalisation of firms: A novel text mining approach* (No. 22-065). ZEW Discussion Papers. https://doi.org/10.2139/ssrn.4321438

Bai, J. J., Brynjolfsson, E., Jin, W., Steffen, S., & Wan, C. (2021). *Digital resilience: How work-from-home feasibility affects firm performance* (No. w28588). National Bureau of Economic Research.

Barney, J. B. (1991). Firm resources and sustained competitive advantage. *Journal of Management, 17*(1), 99-120. https://doi.org/10.1177/014920639101700108

Bunea, O.-I., Corbos, R.-A., & Popescu, R.-I. (2019). Influence of some financial indicators on return on equity ratio in the Romanian energy sector: A competitive approach using a DuPont-based analysis. *Energy*, *189*, 116251. https://doi.org/10.1016/j.energy.2019.116251

Brynjolfsson, E. (1993). The productivity paradox of information technology. *Communications of the ACM*, *36*(12), 66-77. https://doi.org/10.1145/163298.163309

Brynjolfsson, E., & Hitt, L. M. (2000). Beyond computation: Information technology, organizational transformation and business performance. *Journal of Economic Perspectives*, *14*(4), 23-48. https://doi.org/10.1257/jep.14.4.23

Brynjolfsson, E., Horton, J. J., Ozimek, A., Rock, D., Sharma, G., & TuYe, H. Y. (2020). COVID-19 and remote work: An early look at US data (No. w27344). *National Bureau of Economic Research.*

Cai, C., & Zhang, R. (2022). The European Union's digital transformation strategy under the discourse of "technological sovereignty" and "digital sovereignty". *International Political Studies, 43*(1), 9-36.

Calvino, F., Criscuolo, C., Marcolin, L., & Squicciarini, M. (2018). *A taxonomy of digital intensive sectors* (OECD Science, Technology and Industry Working Papers No. 2018/14). OECD Publishing. https://doi.org/10.1787/f404736a-en

Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources, 50*(2), 317-372. https://doi.org/10.3368/jhr.50.2.317

Chen, S.-J., Cheng, C.-J., & Chen, H.-H. (2011). Methodologies for multilingual information integration in the domain of Chinese art. *IFLA Journal, 37*(4), 296-304. https://doi.org/10.1177/0340035211430797

Chen, W., & Srinivasan, S. (2024). Going digital: Implications for firm value and performance. *Review of Accounting Studies, 29*(2), 1619-1665. https://doi.org/10.1007/s11142-023-09753-0

Chen, Y., Cai, Q., Wang, Z., & Xu, Z. (2024). Has digital transformation enhanced the corporate resilience in the face of COVID-19? Evidence from China. *International Review of Financial Analysis*, *96*. https://doi.org/10.1016/j.irfa.2024.103709

China National Committee for Terms in Sciences and Technologies. (n.d.). About TermOnline. *TermOnline*. Retrieved June 11, 2026, from https://www.termonline.cn/about

Ding, W., Levine, R., Lin, C., & Xie, W. (2021). Corporate immunity to the COVID-19 pandemic. *Journal of Financial Economics*, *141*(2), 802-830. https://doi.org/10.1016/j.jfineco.2021.03.005

Gebauer, H., Fleisch, E., Lamprecht, C., & Wortmann, F. (2020). Growth paths for overcoming the digitalization paradox. *Business Horizons*, *63*(3), 313-323. https://doi.org/10.1016/j.bushor.2020.01.005

Gradillas, M., & Thomas, L. D. W. (2023). Distinguishing digitization and digitalization: A systematic review and conceptual framework. *Journal of Product Innovation Management*, *42*(1), 112-143. https://doi.org/10.1111/JPIM.12690

Gu, S., Ma, M. J., Wang, K., Zhang, M., Wei, Z., & Zhou, C. (2025). *How digital transformation influences firm performance: An empirical analysis of Chinese manufacturing firms*. SSRN. https://doi.org/10.2139/ssrn.5111489

Guo, H., Yang, Z., Huang, R., & Guo, A. (2020). The digitalization and public crisis responses of small and medium enterprises: Implications from a COVID-19 survey. *Frontiers of Business Research in China*, *14*(1). https://doi.org/10.1186/s11782-020-00087-1

Guo, L., & Xu, L. (2021). The effects of digital transformation on firm performance: evidence from China's manufacturing sector. *Sustainability (Switzerland)*, *13*(22). https://doi.org/10.3390/su132212844

Hale, T., Angrist, N., Goldszmidt, R., Kira, B., Petherick, A., Phillips, T., Webster, S., Cameron-Blake, E., Hallas, L., Majumdar, S., & Tatlow, H. (2021). A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). *Nature Human Behaviour, 5*, 529-538. https://doi.org/10.1038/s41562-021-01079-8

Härting, R., Gobechia, A., & Tatsumi, I. (2022). Impact of COVID-19 on Digital Business Models—Focus on Sustainability. *Smart Innovation, Systems and Technologies*, *306*, 77-86. https://doi.org/10.1007/978-981-19-3359-2_7

Healy, P. M., & Palepu, K. G. (2001). Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of Accounting and Economics, 31*(1-3), 405-440. https://doi.org/10.1016/S0165-4101(01)00018-0

Helfat, C. E., & Winter, S. G. (2011). Untangling Dynamic and Operational Capabilities: Strategy for the (N)ever-Changing World. *Strategic Management Journal*, *32*(11), 1243-1250. https://doi.org/10.1002/SMJ.955

Herath, T., & Herath, H. S. B. (2020). Coping with the New Normal Imposed by the COVID-19 Pandemic: Lessons for Technology Management and Governance. *Information Systems Management*, *37*(4), 277-283. https://doi.org/10.1080/10580530.2020.1818902

Hotho, A., Nürnberger, A., & Paaß, G. (2005). A brief survey of text mining. *Journal for Language Technology and Computational Linguistics, 20*(1), 19-62.

Hsiao, C. (2007). Panel data analysis—Advantages and challenges. *TEST, 16*(1), 1-22. https://doi.org/10.1007/s11749-007-0046-x

Huang, H., & Zhao, L. (2024). Does digital innovation help firms navigate the COVID-19 pandemic? Evidence from China. *Emerging Markets Review*, *62*. https://doi.org/10.1016/j.ememar.2024.101171

Jardak, M. K., & Ben Hamad, S. (2022). The effect of digital transformation on firm performance: evidence from Swedish listed companies. *The Journal of Risk Finance*, *23*(4), 329-348. https://doi.org/10.1108/JRF-12-2021-0199

Kádárová, J., Lachvajderová, L., & Sukopová, D. (2023). Impact of Digitalization on SME Performance of the EU27: Panel Data Analysis. *Sustainability*, *15*(13), 9973. https://doi.org/10.3390/su15139973

Li, L. (2017). China's manufacturing locus in 2025: With a comparison of "Made-in-China 2025" and "Industry 4.0". *Technological Forecasting and Social Change, 135*, 66-74. https://doi.org/10.1016/j.techfore.2017.05.028

Li, S., Wang, X., & Li, T. (2025). Unveiling the digital tapestry: A review of Chinese and international digital economy research. *Journal of Digital Economy, 4*, 95-107. https://doi.org/10.1016/j.jdec.2025.06.005

Liang, Z., & Wu, P. (2020). International comparison of data governance policies: History, characteristics, and implications. *Science & Technology Review, 38*(5), 36-41.

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

Loughran, T., & McDonald, B. (2016). Textual Analysis in Accounting and Finance: A Survey. *Journal of Accounting Research*, *54*(4), 1187-1230. https://doi.org/10.1111/1475-679X.12123

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance, 66*(1), 35-65. https://doi.org/10.1111/j.1540-6261.2010.01625.x

Mahssouni, R., Makhroute, M., Touijer, M. N., & Elabjani, A. (2023). Dual Perspectives on Financial Performance: Analyzing the Impact of Digital Transformation and COVID-19 on European Listed Companies. *Journal of Risk and Financial Management*, *16*(8). https://doi.org/10.3390/jrfm16080371

Mao, N., & Wang, M. (2024). How does the interaction of digitalization and servitization contribute to manufacturing financial performance? The intermediary effect of manufacturing innovation performance. *Heliyon*, *10*(19). https://doi.org/10.1016/j.heliyon.2024.e38312

Meng, M., Jing, K., & Mander, S. (2017). Scenario analysis of CO_2 emissions from China's electric power industry. *Journal of Cleaner Production, 142*, 3101-3108. https://doi.org/10.1016/j.jclepro.2016.10.157

Nanni, R., Bizzaro, P. G., & Napolitano, M. (2024). The false promise of individual digital sovereignty in Europe: Comparing artificial intelligence and data regulations in China and the European Union. *Policy & Internet, 16,* 711-726. https://doi.org/10.1002/poi3.424

Pletscher-Frankild, S., Pallejà, A., Tsafou, K., Binder, J. X., & Jensen, L. J. (2015). DISEASES: Text mining and data integration of disease-gene associations. *Methods, 74,* 83-89.

Rachinger, M., Rauter, R., Müller, C., Vorraber, W., & Schirgi, E. (2019). Digitalization and its influence on business model innovation. *Journal of Manufacturing Technology Management*, *30*(8), 1143-1160. https://doi.org/10.1108/JMTM-01-2018-0020

Reuschl, A. J., Deist, M. K., & Maalaoui, A. (2022). Digital transformation during a pandemic: Stretching the organizational elasticity. *Journal of Business Research*, *144*, 1320-1332. https://doi.org/10.1016/j.jbusres.2022.01.088

Stavrianou, A., Andritsos, P., & Nicoloyannis, N. (2007). Overview and semantic issues of text mining. *ACM SIGMOD Record, 36(3)*, 23-34.

Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic Capabilities and Strategic Management. Strategic Management Journal, 18(7), 509-533. http://www.jstor.org/stable/3088148

Verrecchia, R. E. (2001). Essays on disclosure. *Journal of Accounting and Economics, 32*(1-3), 97-180. https://doi.org/10.1016/S0165-4101(01)00025-8

Wade, M., & Hulland, J. (2004). Review: The Resource-Based View and Information Systems Research: Review, Extension, and Suggestions For Future Research1.

Management Information Systems Quarterly, *28*(1), 107-142. https://doi.org/10.2307/25148626

Wang, H., & Han, N. (2026). A comparative study of data governance policies between China and the European Union: Based on the PMC index model. *World Sci-Tech R&D, 48*(1), 101-114. https://doi.org/10.16507/j.issn.1006-6055.2025.10.004

Wooldridge, J. M. (2009). Econometrics: Panel data methods. In R. Meyers (Ed.), *Encyclopedia of Complexity and Systems Science*. Springer. https://doi.org/10.1007/978-0-387-30440-3_167

Wu, F., Hu, H., Lin, H., & Ren, X. (2021). Enterprise Digital Transformation and Capital Market Performance Empirical Evidence from Stock Liquidity. *Journal of Management World, 37,* 130-144. https://www.scirp.org/reference/referencespapers?referenceid=3579952

Wu, W., Wang, H., Lu, L., Ma, G., & Gao, X. (2023). How firms cope with social crisis: The mediating role of digital transformation as a strategic response to the COVID-19 pandemic. *Plos one, 18*(4): e0282854. https://doi.org/10.1371/journal.pone.0282854

Yavuz, M. S., Tatlı, H. S., & Bozkurt, G. (2025). Exploring the financial impact of digital transformation: A comprehensive analysis on firms. *Journal of Innovation and Knowledge*, *10*(5). https://doi.org/10.1016/j.jik.2025.100795

Zeng, H., Ran, H., Zhou, Q., Jin, Y., & Cheng, X. (2022). The financial effect of firm digitalization: Evidence from China. *Technological Forecasting and Social Change*, *183*. https://doi.org/10.1016/j.techfore.2022.121951

Zhao, C., Wang, W., & Li, X. (2021). How does digital transformation affect firms' total factor productivity? *Finance & Trade Economics, 42*(7), 114-129. https://doi.org/10.3969/j.issn.1002-8102.2021.07.008

Zheng, K., & Snyder, F. (2023). China and EU's wisdom in choosing competition soft law or hard law in the digital era: A perfect match? *China-EU Law Journal, 9*(1), 25-50. https://doi.org/10.1007/s12689-023-00101-8

Zhou, Y., Ock, Y. S., Alnafrah, I., & Dagestani, A. A. (2023). What aspects explain the relationship between digital transformation and financial performance of firms?. *Journal of Risk and Financial Management*, *16*(11), 479.

Appendices

This section provides detailed explanations and full results for the variable construction process, preliminary model specification tests, supplementary tests, robustness checks, and additional COVID-19 contextual analyses discussed in the main text.

1. Keyword Dictionary and Text Indicator

1.1 Digital Transformation Keyword Dictionary

Appendix Table 10: Digital Transformation Keyword Dictionary

Dimension	Keyword Dictionary
Digital Technology Applications	Data Management, Data Mining, Data Network, Data Platform, Data Center, Data Science, Digital Control, Digital Technology, Digital Communication, Digital Network, Digital Intelligence, Digital Terminal, Digital Marketing, Digitalization, Big Data, Cloud Computing, Cloud IT, Cloud Ecosystem, Cloud Service, Cloud Platform, Blockchain, Internet of Things, Machine Learning
Internet Business Models	Mobile Internet, Industrial Internet, Industry Internet, Internet Solution, Internet Technology, Internet Thinking, Internet Action, Internet Business, Internet Mobile, Internet Application, Internet Marketing, Internet Strategy, Internet Platform, Internet Mode, Internet Business Model, Internet Ecosystem, E-commerce, Electronic Commerce, Internet, Internet+, online/offline, Online to Offline, Online and Offline, O2O, Business-to-Business, Consumer-to-Consumer, Business-to-Consumer, Consumer-to-Business
Intelligent Manufacturing	Artificial Intelligence, High-end Intelligence, Industrial Intelligence, Mobile Intelligence, Intelligent Control, Intelligent Terminal, Intelligent Mobile, Intelligent Management, Intelligent Factory, Intelligent Logistics, Intelligent Manufacturing, Intelligent Warehousing, Intelligent Technology, Intelligent Equipment, Intelligent Production, Intelligent Network, Intelligent System, Intellectualization, Automatic Control, Automatic Monitoring, Automatic Supervision, Automatic Detection, Automatic Production, Numerical Control, Integration, Integrated, Integration Solution, Integrated Control, Integrated System, Industrial Cloud, Future Factory, Intelligent Fault Diagnosis, Life Cycle Management, Manufacturing Execution System, Virtualization, Virtual Manufacturing
Modern Information Systems	Information Sharing, Information Management, Information Integration, Information Software, Information System, Information Network, Information Terminal, Information Center, Informatization, Networked, Industrial Information, Industrial Communication

Source: Based on Zhao *et al.* (2021)

1.2 Translation and Context Check

For the EU listed firm sample, the Chinese keywords need to be translated into English so that they can be searched and identified in English annual reports. To ensure that the English keyword dictionary remains as consistent as possible with the original Chinese keyword dictionary in conceptual meaning, this study constructs the English keyword dictionary through three stages: core term identification, term variant expansion, and semantic verification based on original sentences. This translation and verification process draws on Chen *et al*. (2011), particularly their discussion of concept mapping and contextual adaptation in the construction of Chinese and English multilingual thesauri. Their study points out that the conversion of Chinese and English terms is not always a simple process of literal translation. Terms in different languages may differ because of expression habits, semantic scope, and usage context. Therefore, the construction of a cross language keyword dictionary should focus on the conceptual correspondence between source language terms and target language terms, and should also make appropriate adjustments according to the actual usage context of the target language, rather than relying only on literal translation.

Specifically, in the stage of core term identification, this study first refers to the standardized terms and English equivalents provided by the TermOnline platform[1] developed by the China National Committee for Terms in Sciences and Technologies. TermOnline aims to build a data center, application center, and service center for standardized terminology. It covers standardized terms, new scientific and technological terms, and terminology resources published by the committee over the years, and also incorporates standardized terminology resources released by institutions such as the Central Compilation and Translation Bureau and the China International Communications Group (China National Committee for Terms in Sciences and Technologies, n.d.). Since this platform provides relatively authoritative and standardized Chinese to English term equivalents, it is used as an important reference source for identifying the core English terms in this study.

In the stage of term variant expansion, this study conducts a limited expansion based on the core English terms. The expansion mainly includes capitalization, singular and plural forms, hyphenated forms, common abbreviations, and synonymous or near synonymous expressions that are commonly used in English

[1] TermOnline official website, https://www.termonline.cn/

contexts. For example, electronic commerce and e-commerce refer to the full form and common abbreviated form of the same English term, so they are counted under the same core term during matching. When one Chinese keyword may correspond to several English expressions, this study includes the relevant candidate English terms within the same conceptual scope, provided that their conceptual meaning remains consistent. For example, intelligent factory, smart factory, and intelligent plant may all refer to "intelligent factory" in the context of corporate digitalization, so this study determines whether they should be included in the same category according to the specific context. Similarly, when several Chinese keywords correspond to the same English expression or highly overlapping English concepts, this study merges them to avoid repeated counting and conceptual overlap.

In the stage of semantic verification based on original sentences, this study further checks English terms involving synonymous substitutions, abbreviations, or possible ambiguity. Specifically, the original sentences in which these terms appear in English annual reports are extracted, and the sentence context is used to determine whether their actual meaning remains consistent with the digitalization concept indicated by the original Chinese keyword. If an expanded term shows a clear semantic shift in the context of annual reports, or cannot consistently refer to the concept corresponding to the original Chinese keyword, it is not included in the final English keyword dictionary. Through the above procedure, this study preserves the conceptual framework of the original Chinese keyword system while improving the standardization, coverage, and semantic consistency of English keyword identification as much as possible. Appendix Table 2 presents examples of contextual review from the EU sample, illustrating the semantic judgments used in the keyword screening process.

Appendix Table 11: Contextual Review of English Keywords

Company	Year	Reviewed term	Annual report sentence	Dictionary decision
A2A SpA	2016	Internet of Things / IoT	In 2016 we formed A2A Smart City with the aim of promoting research and development activity to create innovative digital technologies pertaining to the IoT ("Internet of Things"), to be applied to the management of services dedicated to the local community.	Retained
A2A SpA	2020	Artificial Intelligence	Using an artificial intelligence model, it is able to identify and locate leaks, breaks and cracks along the network, enabling predictive maintenance and savings of up to 40% on repairs and upgrades.	Retained

ABN AMRO Bank N.V.	2015	Big Data / Cloud Computing	Technological developments in areas such as mobile banking, social media, data analytics ("big data") and cloud computing create both opportunities and challenges for banks to respond to changing client behaviour and needs.	Retained
Anheuser-Busch InBev SA/NV	2018	Artificial Intelligence / Machine Learning / IoT / Blockchain	Beer Garage is also scaling our existing capabilities in AI & ML, IoT, Cloud & Data Analytics, Automation & Robotics, and exploring emerging technologies such as Blockchain, AR & VR and others.	Retained
Anheuser-Busch InBev SA/NV	2019	Machine Learning	A system that utilizes machine learning to improve and standardize the beer's quality and flavor.	Retained
Anheuser-Busch InBev SA/NV	2018	Data Platform	In 2018, we built a global data platform centralizing all relevant business information, which allowed for deeper analytical insights, driving top-line growth and efficiencies.	Retained
Accor SA	2023	Cloud / Data Platform	This translates to major initiatives such as accelerated migration to the Cloud, updating our hotel management system, strengthening our data platform as well as our central bookings system.	Retained
Acerinox SA	2022	Smart Factory	The Excellence 360 programme includes the Digital Transformation Plan, focused on Industry 4.0, and its path towards the "smart factory".	Retained
Air Liquide SA	2019	Industrial Internet of Things	ST and Air Liquide have developed and tested ergonomic Proofs of Concept which integrate the industrial Internet of Things to improve the traceability of gas cylinders at industrial sites and to provide remote surveillance of connected industrial equipment.	Retained
Ackermans & van Haaren NV	2019	Manufacturing Execution System	The offering in MES (Manufacturing Execution System), a crucial item in the automation of production lines, increased substantially and led to an improvement in market position.	Retained
A2A SpA	2024	DM (Data Mining)	The first is the Ministerial Decree for Incentives aimed at Innovative Renewable Electricity Sources, so-called DM RES II.	Excluded (non-digital meaning)
Aktia Bank Plc	2020	VM (Virtual Manufacturing)	Counterparty risks in derivative instruments are managed through the daily requirement for exchanging securities, ISDA VM Credit Support Annex agreement.	Excluded (non-digital meaning)
Atresmedia Corporación de Medios de Comunicación SA	2015	Autocontrol (Automatic Control)	Atresmedia Televisión is also an active member and part of the founding group of Autorregulación de la Comunicación Comercial, Autocontrol.	Excluded (non-digital meaning)
Accor SA	2020	CNC (Computer Numerical Control)	2005-G of October 12, 2005 of the CNC Emergency Committee on "the circumstances for funding a provision at the parent company of a tax consolidation regime".	Excluded (non-digital meaning)
Acerinox SA	2015	MES (Manufacturing Execution System)	It would be a mistake for the European Union to grant China Market Economy Status (MES), as this would have highly detrimental effects for the industry.	Excluded (ambiguous abbreviation)

Source: Author's elaboration

1.3 Text Processing and Standardization

In terms of text preprocessing, the original annual reports were mostly in PDF format. Direct text extraction from these files may be affected by layout elements such as headers, footers, page numbers, and two column formatting, which can reduce the accuracy of subsequent keyword identification. To improve text processing quality, this study first uses MinerU to convert all annual report PDF files into Markdown format, and then uses Python to read, clean, and organize the converted text in a unified way. This procedure helps reduce the interference of formatting noise in keyword frequency statistics.

For the matching procedure, this study further draws on existing text mining research on term normalization, text preprocessing, and noise control, and applies unified rules to the keyword dictionary and matching process (Hotho *et al*., 2005; Stavrianou *et al*., 2007). For Chinese texts, this study mainly matches core terms and their common expressions according to the existing keyword system. For English texts, in addition to synonymous expressions and common abbreviations, this study also accounts for singular and plural forms as well as British and American spelling differences. For general phrases, the matching process is insensitive to capitalization, spacing, and hyphenation. For special abbreviations that may cause ambiguity, stricter matching rules are applied to reduce the risk of false matches (Pletscher-Frankild *et al*., 2015). At the same time, to avoid overestimating frequency due to repeated expressions, this study uses the sentence as the minimum counting unit. If multiple variants of the same core concept appear in the same sentence, they are counted only once. This procedure produces firm-year level digital keyword frequency data, which provides the basis for constructing the text-based digitalization indicator.

This study denotes the frequency of digitalization related keywords appearing in annual reports as TextDigital, and constructs the text-based digitalization indicator lnTextDigital using ln(TextDigital+1). This indicator captures firms' strategic attention to digitalization issues and the intensity of related disclosure in annual reports. Considering that annual reports differ in length across firms, longer reports may mechanically contain more keywords and thus overstate the degree of digital attention. Therefore, in the robustness tests, this study standardizes keyword frequency by the total word count of annual reports and constructs Std_lnTextDigital as an alternative measure of lnTextDigital.

Although this study makes every effort to control for measurement bias that may arise during cross language identification, the Chinese sample is identified using the original Chinese keyword dictionary, while the EU sample is identified using an English keyword dictionary developed through translation, manual verification, and limited expansion. Therefore, the two samples should be understood as an approximate cross language comparison within a unified research framework, rather than as a completely equivalent measurement across languages.

2. Model Specification Tests

2.1 Multicollinearity Test

Appendix Table 3 reports the VIF results for the explanatory variables used in the main regression models. The mean VIF is 1.12, and all individual VIF values are well below the commonly used threshold of 10, indicating that multicollinearity is unlikely to be a serious concern in the model specification.

Appendix Table 12: Results of VIF Test

Variable	VIF	1/VIF
Size	1.32	0.756174
lnTextDigital	1.23	0.812561
Leverage	1.14	0.877847
Fixed Assets Ratio	1.08	0.921978
ROD2	1.02	0.982892
Cash Flow Ratio	1.01	0.990006
Price to Book Ratio	1.00	0.997558
Mean VIF	1.12	

Source: Author's elaboration

2.2 Panel Unit Root Test

To test the stationarity of the main variables and further reduce the risk of spurious regression, this study conducts panel unit root tests for the core firm level variables. If the null hypothesis of a unit root is rejected, the series can be regarded as stationary (Yavuz *et al.*, 2025). The results in Appendix Table 4 show that ROA, ROE, Tobin's Q, lnTextDigital, ROD2, Size, Leverage, Cash_Flow_Ratio, Fixed_Assets_Ratio, and Price_to_Book_Ratio all pass the test in levels, indicating that these variables are generally stationary. Since COVID_Shock and Stringency_Index only cover the period from 2020 to 2022 and mainly reflect short term external shocks during the pandemic period, they are not included in the panel unit root tests.

Appendix Table 13: Panel Unit Root Test Results

Variable	Level
ROA	8971.72***
ROE	9086.24***
Tobins_Q	16757.20***
Ln(TextDigital+1)	8862.83***

Variable	Level
ROD2	9210.15***
Size	10207.00***
Leverage	9461.85***
Cash_Flow_Ratio	12659.20***
Fixed_Assets_Ratio	10786.10***
Price_to_Book_Ratio	19177.50***

Notes: COVID_Shock and Stringency_Index are excluded because they are only available for the 2020–2022 period. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.
Source: Author's elaboration

2.3 F-test and Hausman Test

This study further compares the panel model specifications using the F-test and the Hausman test. The F-test results are significant for all three dependent variables, indicating that firm specific effects cannot be ignored and that the fixed effects model is preferred to pooled OLS. The Hausman test results are also significant, suggesting that the basic assumptions of the random effects model are not satisfied and that the fixed effects model is more appropriate for the sample used in this study. Therefore, the use of two way fixed effects models in the subsequent analysis is statistically justified.

Appendix Table 14: F-test Results

Dependent variable	Observations	Groups	F-test that all u_i = 0	Prob > F	Conclusion
ROA	25,120	2,512	F(2511, 22592) = 6.11	0.0000	Reject pooled OLS
ROE	25,120	2,512	F(2511, 22592) = 1.90	0.0000	Reject pooled OLS
Tobins_Q	25,120	2,512	F(2511, 22593) = 12.63	0.0000	Reject pooled OLS

Appendix Table 15: Hausman Test Results

Dependent variable	Test statistic	Prob > chi2	Preferred model
ROA	chi2(7) = 818.67	0.0000	Fixed effects
ROE	chi2(7) = 251.97	0.0000	Fixed effects
Tobins_Q	chi2(6) = 419.54	0.0000	Fixed effects

Note: Price-to-Book Ratio is excluded from the Tobin's Q model due to its conceptual overlap with market valuation.
Source: Author's elaboration

2.4 Correlation Matrix

Appendix Table 7 presents the Pearson correlation matrix for the variables. The results show that although some pairwise correlations are statistically significant, most coefficients are small in absolute value, and the variables do not show widespread high correlations. The correlation between COVID_Shock and Stringency_Index is -0.690, and that between ROE and Price_to_Book_Ratio is -0.628. These values are relatively high, but they remain below the level commonly considered to raise serious multicollinearity concerns.

Appendix Table 16: Pearson Correlation Matrix

Variable	ROA	ROE	Tobins_Q	lnTextDigital	ROD2	COVID_Shock	Stringency_Index	Size	Leverage	Cash_Flow_Ratio	Fixed_Assets_Ratio	Price_to_Book_Ratio
ROA	1.000											
ROE	0.331***	1.000										
Tobins_Q	0.143***	0.003	1.000									
lnTextDigital	-0.005	0.008	-0.009	1.000								
ROD2	0.104***	0.012*	0.011*	0.024***	1.000							
COVID_Shock	0.013	0.055***	-0.102***	-0.072***	0.044***	1.000						
Stringency_Index	-0.025**	-0.044***	0.022*	0.121***	-0.073***	-0.690***	1.000					
Size	0.088***	0.054***	-0.187***	0.343***	0.125***	0.011	0.044***	1.000				
Leverage	-0.268***	-0.087***	-0.291***	0.048***	0.055***	0.218***	-0.112***	0.330***	1.000			
Cash_Flow_Ratio	0.144***	0.036***	0.025***	0.001	0.036***	-0.002	0.007	0.043***	-0.049***	1.000		
Fixed_Assets_Ratio	0.014**	-0.003	-0.100***	-0.242***	0.007	0.085***	-0.092***	0.027***	-0.029***	0.054***	1.000	
Price_to_Book_Ratio	-0.021***	-0.628***	0.184***	-0.005	0.001	-0.051***	-0.003	-0.037***	0.017***	-0.004	-0.005	1.000

Notes: *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.
Source: Author’s elaboration

3. Supplementary and Robustness Tests

3.1 ROD2 Supplementary Tests

3.1.1 Association Between lnTextDigital and ROD2

To preliminarily examine whether there is an empirical association between ROD2 and the text based digitalization indicator, this study uses ROD2 as the dependent variable and lnTextDigital as the explanatory variable, and further conducts a test using one period lagged lnTextDigital. The results show a positive and weakly significant association between lnTextDigital and ROD2. This suggests that firms with stronger textual digital attention also tend to display higher asset output efficiency, although the association may also reflect common firm characteristics such as size, industry conditions, management quality, or overall operating efficiency. Therefore, the result provides suggestive support for using ROD2 as a complementary operational indicator, rather than evidence that ROD2 is a digital specific measure.

Appendix Table 17: Association Between lnTextDigital and ROD2

Variables	Current lnTextDigital	Lagged lnTextDigital
	ROD2	ROD2
lnTextDigital	0.0076*	
	(0.0045)	
L.lnTextDigital		0.0078*
		(0.0043)
Size	0.0382***	0.0454***
	(0.0089)	(0.0094)
Leverage	-0.0292	-0.0503
	(0.0369)	(0.0365)
Cash Flow Ratio	0.0157	0.0235
	(0.0115)	(0.0207)
Fixed Assets Ratio	0.0296	0.0262
	(0.0512)	(0.0570)
Observations	25,120	22,608
Adjusted R^2	0.8365	0.8533

Notes: Robust standard errors clustered at the firm level are reported in parentheses. Firm and year fixed effects are included. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.
Source: Author's elaboration.

3.1.2 High and Low Digital-Intensity Industry Test

To examine industry differences in the performance associations of the digitalization indicators, this study classifies sample firms according to the Organization for Economic Co-operation and Development (OECD) industry digital intensity classification proposed by Calvino et al. (2018), based on International Standard Industrial Classification (ISIC) Rev. 4 industry codes. This classification considers multiple dimensions, including ICT investment, ICT intermediate inputs, robot use, ICT specialists, and online sales, and divides industries into four categories: High, Medium high, Medium low, and Low. This study further combines High and Medium high into high digital intensity industries, and Medium low and Low into low digital intensity industries. A dummy variable, Highlow, is then constructed. It equals 1 if a firm belongs to a high digital intensity industry, and 0 otherwise.

The results are reported in Appendix Table 9. lnTextDigital has a significantly positive effect on ROA in both high and low digital intensity industries, and also shows a weakly significant positive effect on ROE in low digital intensity industries. However, its effect on Tobin's Q is not significant in either group. By contrast, ROD2 has a significantly positive effect on ROA, ROE, and Tobin's Q in high digital intensity industries, while in low digital intensity industries it is significant only for Tobin's Q. This indicates that the positive relationship between ROD2 and accounting performance is mainly concentrated in high digital intensity industries.

Since differences in significance across subsamples do not necessarily imply significant differences between coefficients, this study adds interaction terms between the digitalization variables and Highlow in the full sample. The results are reported in Appendix Table 10. The effect of lnTextDigital × Highlow is not significant for ROA or ROE, but is significantly positive for Tobin's Q, suggesting that the marginal relationship between textual digital attention and market valuation is relatively stronger in high digital intensity industries. ROD2 × Highlow is significantly positive for ROA and ROE at the 10% and 5% levels, respectively, but its effect on Tobin's Q is not significant.

Overall, the relationship between ROD2 and firm performance is more evident in high digital intensity industries, especially for accounting performance indicators such as ROA and ROE. This result suggests that the relationship between ROD2 and accounting performance is more pronounced in high digital-intensity industries, where digital infrastructure and technology use may be more developed. Therefore, the industry subsample test provides supplementary evidence for interpreting ROD2 from the perspective of operational conversion. At the same time, the meaning of ROD2 should still be limited to asset output efficiency, and its digitalization related interpretation should be understood together with the text indicator, industry context, and other robustness tests.

Appendix Table 18: High and Low Digital-Intensity Industry Subsample Results

Panel A: lnTextDigital						
Variables	ROA	ROE	Tobin's Q	ROA	ROE	Tobin's Q
	High Digital-Intensity Industries			Low Digital-Intensity Industries		
lnTextDigital	0.4730**	1.3116	-0.0102	0.2834***	2.2705*	-0.0186
	(0.2148)	(0.9198)	(0.0310)	(0.1025)	(1.1876)	(0.0205)
Size	-1.3038	2.4344	-0.0047	1.4953***	5.3810***	-0.2403***
	(2.3086)	(4.4578)	(0.0618)	(0.2148)	(1.7112)	(0.0435)
Leverage	-24.5215***	-45.0039***	0.1134	-19.9623***	-71.9796***	0.0420
	(3.4863)	(12.3767)	(0.2874)	(1.2454)	(9.5805)	(0.1784)
Cash Flow Ratio	1.3502	1.4978	-0.0260	1.0569	1.7020	0.0284
	(1.5479)	(1.6421)	(0.0252)	(1.0702)	(2.0478)	(0.0318)
Fixed Assets Ratio	-8.2906**	-47.6175***	0.7900**	-11.4834***	-5.4168	0.1411
	(3.4354)	(13.1993)	(0.3585)	(1.1171)	(20.8749)	(0.3263)
Price-to-Book Ratio	-0.0818	3.8011		0.0013	-1.4734***	
	(0.3104)	(2.8134)		(0.0017)	(0.1244)	
Constant	27.7106	8.8050	1.4735***	2.1831	-7.4234	3.8479***
	(17.9280)	(38.2826)	(0.4953)	(1.6928)	(8.5504)	(0.3173)
Observations	4,690	4,690	4,690	20,430	20,430	20,430
Adjusted R^2	0.4716	0.1483	0.6770	0.3978	0.5170	0.5968
Panel B: ROD2						
Variables	ROA	ROE	Tobin's Q	ROA	ROE	Tobin's Q
	High Digital-Intensity Industries			Low Digital-Intensity Industries		
ROD2	6.7527***	15.7206***	0.6185***	1.8324	-2.8226	0.3552***
	(2.2627)	(5.4712)	(0.2048)	(1.3755)	(5.4603)	(0.0567)
Size	-1.4267	2.1935	-0.0263	1.4989***	6.1476***	-0.2605***

	(2.2779)	(4.3986)	(0.0607)	(0.2187)	(1.5565)	(0.0433)
Leverage	-24.7280***	-45.5223***	0.0807	-19.8817***	-72.3656***	0.0643
	(3.4457)	(12.2892)	(0.2861)	(1.2631)	(9.6761)	(0.1781)
Cash Flow Ratio	1.3289	1.4521	-0.0289	1.0115	1.7178	0.0210
	(1.4955)	(1.5241)	(0.0242)	(1.0316)	(2.0925)	(0.0252)
Fixed Assets Ratio	-7.9411**	-47.0043***	0.8674**	-11.7633***	-6.5614	0.1296
	(3.8452)	(13.8956)	(0.3601)	(1.1037)	(20.2580)	(0.3240)
Price-to-Book Ratio	-0.1241	3.7024		0.0012	-1.4734***	
	(0.3080)	(2.8147)		(0.0017)	(0.1244)	
Constant	26.0163	5.2145	1.2241***	1.9880	-4.1319	3.7286***
	(17.4228)	(37.3308)	(0.4731)	(1.7188)	(9.1677)	(0.3181)
Observations	4,690	4,690	4,690	20,430	20,430	20,430
Adjusted R^2	0.4788	0.1513	0.6848	0.4004	0.5167	0.5992

Notes: Robust standard errors clustered at the firm level are reported in parentheses. Firm and year fixed effects are included. Price-to-Book Ratio is excluded from the Tobin's Q models. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.
Source: Author's elaboration.

Appendix Table 19: Industry Digital Intensity Interaction Results

Variables	ROA	ROE	Tobin's Q	ROA	ROE	Tobin's Q
	Panel A: lnTextDigital × Highlow			Panel B: ROD2 × Highlow		
lnTextDigital	0.2907***	1.6894	-0.0511**			
	(0.1079)	(1.1612)	(0.0202)			
lnTextDigital × Highlow	0.2682	1.8215	0.1678***			
	(0.2168)	(1.4094)	(0.0353)			
ROD2				1.8725	-2.7472	0.3438***
				(1.3748)	(5.4759)	(0.0552)
ROD2 × Highlow				3.9074*	23.7962**	0.1990
				(2.3297)	(9.6061)	(0.1931)
Size	0.9030	4.0412**	-0.1909***	0.8850	4.4869***	-0.2148***
	(0.6387)	(1.7067)	(0.0376)	(0.6329)	(1.5980)	(0.0378)
Leverage	-21.0503***	-57.9632***	0.0491	-21.0880***	-58.6404***	0.0513
	(1.3883)	(10.7928)	(0.1534)	(1.3920)	(10.7087)	(0.1525)
Cash Flow Ratio	1.1803	1.3386	-0.0009	1.1405	1.3186	-0.0069
	(0.9051)	(1.3318)	(0.0236)	(0.8765)	(1.3113)	(0.0218)
Fixed Assets Ratio	-11.1720***	-8.1956	0.2674	-11.2992***	-8.5677	0.2996
	(1.0568)	(18.4469)	(0.2897)	(1.0605)	(17.8058)	(0.2872)
Price-to-Book Ratio	0.0012	-1.4651***		0.0010	-1.4653***	
	(0.0018)	(0.1272)		(0.0018)	(0.1272)	
Constant	7.3935	-0.0138	3.3763***	7.0589	2.1034	3.2615***

	(4.7902)	(11.3425)	(0.2781)	(4.7649)	(11.7533)	(0.2807)
Observations	25,120	25,120	25,120	25,120	25,120	25,120
Adjusted R^2	0.4254	0.4562	0.6048	0.4291	0.4567	0.6068

Notes: Robust standard errors clustered at the firm level are reported in parentheses. Firm and year fixed effects are included. Price-to-Book Ratio is excluded from the Tobin's Q models. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.
Source: Author's elaboration.

3.1.3 Interaction Between lnTextDigital and ROD2

To further examine whether there is a complementary relationship between textual digital attention and the operational conversion path reflected by ROD2, this study includes lnTextDigital, ROD2, and their interaction term in the same model. The results show that, after controlling for ROD2, lnTextDigital still has a significantly positive effect on ROA. After controlling for lnTextDigital, ROD2 still has a significantly positive effect on Tobin's Q.

The coefficients of lnTextDigital × ROD2 for ROA, ROE, and Tobin's Q are not statistically significant. This result suggests that textual digital attention and the operational conversion path reflected by ROD2 do not show a significant complementary relationship in the sample. Therefore, the two indicators are better interpreted as related but relatively independent observation dimensions, rather than as a direct sequence from textual digital strategy to operational realization.

Appendix Table 20: Interaction Between lnTextDigital and ROD2

Variables	ROA	ROE	Tobin's Q
lnTextDigital	0.5385***	1.0653	-0.0323
	(0.1895)	(1.1078)	(0.0241)
ROD2	3.5724*	-5.4556	0.3161***
	(1.9467)	(10.3234)	(0.1099)
lnTextDigital × ROD2	-0.3631	1.6101	0.0154
	(0.2907)	(2.6448)	(0.0267)
Size	0.7966	4.0608**	-0.2075***
	(0.6355)	(1.6724)	(0.0381)
Leverage	-20.9599***	-58.2034***	0.0500
	(1.4011)	(10.8534)	(0.1525)
Cash Flow Ratio	1.1378	1.3596	-0.0067
	(0.8764)	(1.3426)	(0.0218)
Fixed Assets Ratio	-11.2110***	-7.9261	0.2792
	(1.0618)	(18.3382)	(0.2891)

Price-to-Book Ratio	0.0011	-1.4651***	
	(0.0018)	(0.1271)	
Constant	6.1127	3.1173	3.3219***
	(4.7915)	(13.2261)	(0.2905)
Observations	25,120	25,120	25,120
Adjusted R^2	0.4287	0.4563	0.6068

Notes: Robust standard errors clustered at the firm level are reported in parentheses. Firm and year fixed effects are included. Price-to-Book Ratio is excluded from the Tobin's Q models. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.
Source: Author's elaboration.

3.1.4 Cash Flow Ratio as an Alternative Dependent Variable

To further address the possible mechanical overlap between ROD2 and accounting performance indicators, this study uses Cash Flow Ratio as an alternative dependent variable for the robustness test. Compared with ROA and ROE, Cash Flow Ratio reflects firm performance from the perspective of operating cash flow, rather than measuring profitability directly based on accounting profit. In this model, Cash Flow Ratio is measured as operating cash flow divided by total assets, and therefore it is no longer included as a control variable.

As shown in Appendix Table 12, ROD2 remains significantly positively associated with Cash Flow Ratio. The coefficient of ROD2 is 0.0337 and is significant at the 1% level. This result suggests that the positive relationship between ROD2 and firm performance does not depend entirely on accounting profit based performance indicators such as ROA and ROE. Therefore, this robustness test provides further support for interpreting ROD2 as an indicator related to asset output efficiency and operational conversion.

Appendix Table 21: Cash Flow Ratio as an Alternative Dependent Variable

Variables	Cash Flow Ratio
ROD2	0.0337***
	(0.0066)
Size	0.0001
	(0.0050)
Leverage	-0.0915***
	(0.0185)
Fixed Assets Ratio	0.0009
	(0.0177)
Price-to-Book Ratio	-0.0000**

	(0.0000)
Constant	0.0769*
	(0.0424)
Observations	25,120
Adjusted R^2	0.0336

Notes: Robust standard errors clustered at the firm level are reported in parentheses. Firm and year fixed effects are included. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.
Source: Author's elaboration

3.1.5 Additional Control for Profit Margin

In addition, this study further adds Profit Margin as an additional control variable to the baseline model. After controlling for Profit Margin, ROD2 remains positive and significant for ROA and Tobin's Q, while its effect on ROE remains insignificant, consistent with the baseline regression results. This indicates that the main results for ROD2 do not change substantially after controlling for the profit margin component. Therefore, the relationship between ROD2 and firm performance is not fully explained by profit margin.

Appendix Table 22: Additional Test Controlling for Profit Margin

Variables	ROA	ROE	Tobin's Q
ROD2	2.2800*	0.0496	0.3687***
	(1.2497)	(5.1291)	(0.0547)
Size	1.5566***	5.6162***	-0.2269***
	(0.2122)	(1.4379)	(0.0384)
Leverage	-21.0840***	-60.4682***	0.0473
	(1.2634)	(10.8206)	(0.1525)
Cash Flow Ratio	1.0022	1.5663	0.0010
	(0.7775)	(1.4525)	(0.0226)
Fixed Assets Ratio	-11.4750***	-9.3149	0.3039
	(1.0444)	(17.8080)	(0.2872)
Price-to-Book Ratio	0.0018	-1.4687***	
	(0.0016)	(0.1262)	
Profit Margin	0.0000	-0.0000	0.0000**
	(0.0000)	(0.0000)	(0.0000)
Constant	2.0320	-4.7024	3.3670***
	(1.6111)	(10.3478)	(0.2853)
Observations	25090	25090	25090
Adjusted R^2	0.4268	0.4807	0.6081

Notes: Robust standard errors clustered at the firm level are reported in parentheses. Firm and year fixed effects are included. Price-to-Book Ratio is excluded from the Tobin's Q models. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.
Source: Author's elaboration

3.2 Alternative Measures

3.2.1 Standardized Text Indicator

To examine whether the text based digitalization indicator is affected by differences in annual report length, this study re-estimates the baseline regressions using lnTextDigital standardized by the total word count of annual reports.

Appendix Table 14 reports the robustness test results based on this standardized text indicator. The results show that Std_lnTextDigital still has a significantly positive effect on ROA and ROE, while its effect on Tobin's Q is not significant. Compared with the baseline regression results, the positive relationship between the text based digitalization indicator and accounting performance remains stable after standardization. This suggests that the conclusion regarding the positive relationship between textual digital attention and accounting performance is not simply driven by differences in annual report length.

Appendix Table 23: Robustness Test Based on Std_lnTextDigital

Variables	ROA	ROE	Tobin's Q
Std_lnTextDigital	0.4079***	2.4280**	-0.0244
	(0.1144)	(1.2001)	(0.0215)
Size	0.8981	4.0083**	-0.1940***
	(0.6366)	(1.6907)	(0.0378)
Leverage	-21.0642***	-58.0572***	0.0406
	(1.3913)	(10.7711)	(0.1535)
Cash Flow Ratio	1.1799	1.3362	-0.0012
	(0.9051)	(1.3316)	(0.0239)
Fixed Assets Ratio	-11.1362***	-7.9523	0.2899
	(1.0527)	(18.3367)	(0.2888)
Price-to-Book Ratio	0.0012	-1.4650***	
	(0.0018)	(0.1272)	
Constant	8.5150*	6.6940	3.3355***
	(4.8331)	(12.1082)	(0.2811)
Observations	25,120	25,120	25,120
Adjusted R^2	0.4253	0.4562	0.6041

Notes: Robust standard errors clustered at the firm level are reported in parentheses. Firm and year fixed effects are included. Price-to-Book Ratio is excluded from the Tobin's Q models. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Source: Author’s elaboration

3.2.2 Sensitivity Test Excluding Broad and Context-Dependent Terms

To further assess whether the baseline results are driven by broad and context dependent keywords, this study reconstructs lnTextDigital by excluding selected generic terms, including integration, information system, networked, information management, data management, information sharing, information network, information integration, intelligent technology, industrial information and life cycle management. The results remain broadly consistent with the baseline pattern. The reconstructed lnTextDigital remains positively associated with ROA and ROE, while its association with Tobin’s Q remains statistically insignificant. This suggests that the accounting performance pattern is not solely driven by generic informatization terms, although the smaller coefficients indicate that the text based indicator is sensitive to dictionary scope.

Appendix Table 24: Sensitivity Test Excluding Broad Digital Terms

Variables	ROA	ROE	Tobin's Q
Strict lnTextDigital	0.1434*	1.4221*	0.0027
	(0.0810)	(0.7522)	(0.0151)
Original lnTextDigital	0.3307***	1.9533*	-0.0254
	(0.1060)	(1.0288)	(0.0183)
Size	0.8548	4.1465**	-0.2151***
	(0.6289)	(1.6960)	(0.0382)
Leverage	-21.0135***	-58.1199***	0.0545
	(1.4001)	(10.8086)	(0.1525)
Cash Flow Ratio	1.1439	1.3467	-0.0068
	(0.8801)	(1.3334)	(0.0220)
Fixed Assets Ratio	-11.3146***	-8.2881	0.2957
	(1.0564)	(18.2241)	(0.2889)
Price-to-Book Ratio	0.0010	-1.4650***	
	(0.0018)	(0.1274)	
Constant	7.0880	2.0297	3.2651***
	(4.7726)	(11.6116)	(0.2815)
Observations	25,120	25,120	25,120
Adjusted R^2	0.4282	0.4561	0.6067

Notes: Robust standard errors clustered at the firm level are reported in parentheses. Firm and year fixed effects are included. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.
Source: Author’s elaboration.

3.2.3 Alternative Operational Conversion Indicator

To further examine whether the construction of the operational conversion indicator affects the results, this study uses lnROD to replace ROD2 in the robustness test. lnROD is constructed from the ratio of operating revenue to the number of employees, and reflects operating efficiency from the perspective of unit labor output.

The results show that lnROD has a significantly positive effect on both ROA and ROE, indicating that operating efficiency measured by unit labor output can also explain differences in accounting performance. This result is broadly consistent with the main analysis, where the operating efficiency dimension is positively associated with accounting performance. It suggests that, after replacing ROD2 with a labor output efficiency indicator, the positive relationship between operating efficiency and accounting performance still holds. However, the coefficient of lnROD on Tobin's Q is -0.1140 and is significantly negative at the 5% level, which differs from the significantly positive effect of ROD2 on Tobin's Q in the main analysis. This difference may arise from the different efficiency dimensions captured by the two indicators. ROD2 mainly reflects asset output efficiency, while ROD mainly reflects labor output efficiency, so market valuation may not respond to the two efficiency dimensions in the same way.

Appendix Table 25: Robustness Test Based on lnROD

Variables	ROA	ROE	Tobin's Q
lnROD	3.7618***	9.0375***	-0.1140**
	(0.3655)	(1.6111)	(0.0443)
Size	2.5223***	7.7137***	-0.2379***
	(0.1967)	(1.4366)	(0.0402)
Leverage	-21.7638***	-61.9286***	0.0546
	(1.0844)	(11.1994)	(0.1508)
Cash Flow Ratio	0.9161	1.2783	0.0105
	(0.7072)	(1.2244)	(0.0272)
Fixed Assets Ratio	-9.4480***	-4.5979	0.2535
	(1.0792)	(18.2414)	(0.2946)
Price-to-Book Ratio	0.0036**	-1.4649***	
	(0.0017)	(0.1275)	
Constant	-55.8371***	-145.3039***	5.2455***
	(5.8274)	(26.4347)	(0.7838)
Observations	25,089	25,089	25,089
Adjusted R^2	0.4501	0.4842	0.6061

Notes: Robust standard errors clustered at the firm level are reported in parentheses. Firm and year fixed effects are included. Price-to-Book Ratio is excluded from the Tobin's Q models. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.
Source: Author's elaboration

3.3 Sample and Model Checks

3.3.1 Winsorization

To further test whether the baseline regression results are affected by extreme observations, this study winsorizes the main continuous variables at the 1st and 99th percentiles and repeats the H1 baseline model using the winsorized sample. Appendix Table 17 reports the corresponding regression results.

Panel A shows that, after winsorization, lnTextDigital still has a significantly positive effect on ROA and ROE, with coefficients of 0.3906 and 1.2413, respectively, both significant at the 1% level. At the same time, the effect of lnTextDigital on Tobin's Q remains statistically insignificant. This result is broadly consistent with the baseline regression results, indicating that the positive relationship between textual digital attention and accounting performance is not driven by a small number of extreme observations.

Panel B shows that, after winsorization, ROD2 has a significantly positive effect on ROA, ROE, and Tobin's Q, with coefficients of 3.7535, 8.0827, and 0.4667, respectively, all significant at the 1% level. Compared with the baseline regression, the significantly positive relationship between ROD2 and ROA and Tobin's Q remains unchanged. The positive effect on ROE reaches statistical significance after winsorization.

Overall, the winsorized results are broadly consistent with the baseline findings. This suggests that the main conclusions are not driven by a small number of extreme observations.

Appendix Table 26: Robustness Test Based on Winsorized Variables

Variables	ROA	ROE	Tobin's Q	ROA	ROE	Tobin's Q
	Panel A: lnTextDigital			Panel B: ROD2		
lnTextDigital	0.3906***	1.2413***	-0.0153			
	(0.0696)	(0.2023)	(0.0133)			
ROD2				3.7535***	8.0827***	0.4667***
				(0.3364)	(0.9400)	(0.0516)
Size	1.3023***	3.7200***	-0.1896***	1.2065***	3.6203***	-0.2147***
	(0.2184)	(0.6075)	(0.0291)	(0.2126)	(0.5949)	(0.0295)
Leverage	-16.8362***	-39.4283***	-0.0293	-16.5675***	-38.8912***	-0.0088

	(0.6499)	(2.1779)	(0.1166)	(0.6500)	(2.1778)	(0.1160)
Cash Flow Ratio	23.4770***	48.9125***	1.1769***	21.5076***	44.6511***	0.9236***
	(1.0427)	(2.6270)	(0.1265)	(0.9989)	(2.5880)	(0.1212)
Fixed Assets Ratio	-10.2786***	-20.5873***	0.5179***	-10.5174***	-21.3585***	0.5269***
	(0.7442)	(2.0995)	(0.1522)	(0.7647)	(2.1376)	(0.1545)
Price-to-Book Ratio	0.3937***	0.3009***		0.3537***	0.2133**	
	(0.0289)	(0.1036)		(0.0291)	(0.1027)	
Constant	-0.4592	-8.2898*	3.2375***	-0.5500	-7.9813*	3.1117***
	(1.6732)	(4.6277)	(0.2235)	(1.6434)	(4.5741)	(0.2248)
Observations	25,120	25,120	25,120	25,120	25,120	25,120
Adjusted R^2	0.5514	0.4289	0.6646	0.5597	0.4346	0.6685

Notes: Variables are winsorized at the 1st and 99th percentiles. Robust standard errors clustered at the firm level are reported in parentheses. Firm and year fixed effects are included. Price-to-Book Ratio is excluded from the Tobin's Q models. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.
Source: Author's elaboration

3.3.2 Sector-Year Fixed Effects

To further control for common technological, policy, and macroeconomic shocks faced by firms in the same industry in a given year, this study adds ISIC industry × year fixed effects on the basis of firm fixed effects and re-estimates the baseline models. The results show that, after controlling for industry × year fixed effects, the effects of lnTextDigital on ROA and ROE remain significantly positive at the 1% and 5% levels, respectively, while its effect on Tobin's Q remains insignificant. This result indicates that the positive relationship between textual digitalization attention and current accounting performance is not fully driven by industry level annual common shocks.

For ROD2, its effect on Tobin's Q remains significantly positive at the 1% level, but its coefficients on ROA and ROE do not reach statistical significance. Overall, after adding industry × year fixed effects, the results for the text based digitalization indicator on accounting performance and the result for ROD2 on market valuation remain valid. However, the effect of ROD2 on accounting performance shows some sensitivity and may be affected by changes in the industry year environment.

Appendix Table 27: Robustness Test with Sector-Year Fixed Effects

Variables	ROA	ROE	Tobin's Q	ROA	ROE	Tobin's Q
	Panel A: lnTextDigital			Panel B: ROD2		
lnTextDigital	0.2527***	2.0280**	-0.0128			

	(0.0978)	(0.9661)	(0.0175)			
ROD2				1.8733	-1.8556	0.3539***
				(1.3214)	(5.3625)	(0.0590)
Size	0.8878	4.0935**	-0.1868***	0.8876	4.6723***	-0.2021***
	(0.6085)	(1.7601)	(0.0382)	(0.6042)	(1.6759)	(0.0381)
Leverage	-20.7898***	-58.3902***	0.1027	-20.8137***	-58.5635***	0.1029
	(1.3583)	(10.7552)	(0.1449)	(1.3641)	(10.7772)	(0.1445)
Cash Flow Ratio	1.0993	1.2026	-0.0090	1.0722	1.2159	-0.0138
	(0.8194)	(1.2803)	(0.0269)	(0.8006)	(1.2944)	(0.0255)
Fixed Assets Ratio	-10.5393***	-8.2595	0.4019	-10.7727***	-9.2940	0.3913
	(1.0776)	(19.2728)	(0.2699)	(1.0742)	(18.8310)	(0.2702)
Price-to-Book Ratio	0.0012	-1.4658***		0.0010	-1.4659***	
	(0.0020)	(0.1268)		(0.0020)	(0.1268)	
Constant	7.5216*	-0.2672	3.2632***	7.2514	3.0161	3.1319***
	(4.5424)	(12.1259)	(0.2894)	(4.5643)	(12.6923)	(0.2907)
Observations	25060	25060	25060	25,060	25,060	25,060
Adjusted R^2	0.4396	0.4658	0.6286	0.4414	0.4655	0.6309

Notes: Robust standard errors clustered at the firm level are reported in parentheses. Firm and ISIC industry × year fixed effects are included. Price-to-Book Ratio is excluded from the Tobin's Q models. Six firms belonging to singleton ISIC sectors, corresponding to 60 firm year observations, are omitted when absorbing firm and sector year fixed effects. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.
Source: Author's elaboration.

3.3.3 Lead Variable Test

To examine the time direction between the core variables and firm performance, this study re-estimates the models using one period lead values of lnTextDigital and ROD2. If future digitalization variables are still significantly associated with current performance, this may suggest the presence of persistence or reverse association, meaning that firms with higher performance may also show stronger digital strategic attention or higher asset output efficiency in subsequent periods.

The results are reported in Appendix Table 19. Panel A shows that F1.lnTextDigital is significantly positive for ROA and ROE, but insignificant for Tobin's Q. Panel B shows that F1.ROD2 is significantly positive for ROA and Tobin's Q, but insignificant for ROE. These results suggest that there is an empirical association between firm performance and subsequent digital strategic attention and asset output efficiency. Therefore, the main regression results in this study are more appropriately interpreted as associations rather than strict causal relationships. At the same time, the results for F1.ROD2 also suggest that firms with higher current performance may also be associated with higher asset output efficiency in subsequent periods. Therefore, the interpretation of ROD2 should still be understood together with its asset output efficiency attribute and the other robustness tests.

Appendix Table 28: Lead Variable Test Results

Variables	ROA	ROE	Tobin's Q	ROA	ROE	Tobin's Q
	Panel A: F1.lnTextDigital			Panel B: F1.ROD2		
F1.lnTextDigital	0.6000***	1.3286***	0.0017			
	(0.1056)	(0.5107)	(0.0192)			
F1.ROD2				1.6599***	5.1968	0.1751***
				(0.4579)	(5.0859)	(0.0480)
Size	1.1093*	3.2943**	-0.1648***	1.1846*	3.4284**	-0.1682***
	(0.6503)	(1.6337)	(0.0414)	(0.6521)	(1.6368)	(0.0413)
Leverage	-24.0028***	-50.9476***	-0.1163	-24.2895***	-51.7370***	-0.1344
	(2.7987)	(10.9014)	(0.1720)	(2.7937)	(10.9034)	(0.1726)
Cash Flow Ratio	0.9705	0.6103	0.0030	0.9546	0.5570	0.0010
	(0.7688)	(0.9830)	(0.0236)	(0.7634)	(0.9728)	(0.0229)
Fixed Assets Ratio	-9.8182***	1.8837	0.3494	-10.1808***	1.0504	0.3450
	(0.9976)	(20.4401)	(0.3298)	(0.9912)	(20.2213)	(0.3280)
Price-to-Book Ratio	0.0003	-0.8966***		0.0003	-0.8966***	
	(0.0082)	(0.2380)		(0.0082)	(0.2383)	
Constant	6.1207	1.4184	3.1819***	6.6784	2.0692	3.1181***
	(4.5941)	(11.4468)	(0.3104)	(4.6284)	(11.7511)	(0.3053)
Observations	22,608	22,608	22,608	22,608	22,608	22,608
Adjusted R^2	0.4266	0.1112	0.6177	0.4269	0.1115	0.6182

Notes: Robust standard errors clustered at the firm level are reported in parentheses. Firm and year fixed effects are included. Price-to-Book Ratio is excluded from the Tobin's Q models. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.
Source: Author's elaboration

4. Additional COVID-19 Analysis

4.1 Country-level Stringency Indicator

In the analysis of policy stringency during the pandemic, Stringency_during is matched at the provincial level for the Chinese sample and at the country level for the EU sample. Although this setting can more accurately capture policy differences across Chinese regions, it may also create an inconsistency in the geographic matching level between the Chinese and EU samples. To further test whether this issue affects the conclusions of this study, this study uses a unified country level policy stringency indicator, CountryStringency_during, to re-estimate the models.

The results in Appendix Table 20 are broadly consistent with the provincial level results. After using the country level policy stringency indicator for all samples, lnTextDigital × CountryStringency_during is not significant for ROA or ROE, but shows a weakly significant negative effect on Tobin's Q. ROD2 × CountryStringency_during remains significantly positive for ROA and ROE, while its effect on Tobin's Q is not significant. This result is broadly consistent with the analysis based on the original policy stringency indicator in the main text, indicating that the main estimation results do not change substantially after policy stringency is measured at the country level.

Appendix Table 29: Country-level Policy Stringency Results

Variables	ROA	ROE	Tobin's Q	ROA	ROE	Tobin's Q
	Panel A: lnTextDigital × policy stringency			Panel B: ROD2 × policy stringency		
lnTextDigital	0.3756***	2.1812**	-0.0114			
	(0.0950)	(1.0338)	(0.0191)			
ROD2				2.0520	-1.7148	0.3641***
				(1.3233)	(5.2924)	(0.0565)
CountryStringency_during	-0.0074	-0.0917	-0.0021	-0.0228**	-0.1539*	-0.0043***
	(0.0107)	(0.0911)	(0.0013)	(0.0109)	(0.0848)	(0.0008)
lnTextDigital × CountryStringency_during	-0.0018	-0.0053	-0.0004*			
	(0.0016)	(0.0065)	(0.0003)			
ROD2 × CountryStringency_during				0.0149***	0.0941**	0.0004
				(0.0055)	(0.0367)	(0.0005)
Size	0.9074	4.0648**	-0.1916***	0.9031	4.6039***	-0.2119***
	(0.6354)	(1.6893)	(0.0378)	(0.6327)	(1.5980)	(0.0378)
Leverage	-21.0407***	-57.9829***	0.0462	-21.0494***	-58.4098***	0.0536
	(1.3944)	(10.7700)	(0.1532)	(1.3997)	(10.8446)	(0.1525)

Cash Flow Ratio	1.1809	1.3409	-0.0009	1.1403	1.3168	-0.0069
	(0.9072)	(1.3450)	(0.0238)	(0.8814)	(1.3415)	(0.0217)
Fixed Assets Ratio	-11.1876***	-8.3632	0.2764	-11.4527***	-9.5521	0.2770
	(1.0489)	(18.3695)	(0.2894)	(1.0439)	(17.8198)	(0.2871)
Price-to-Book Ratio	0.0012	-1.4653***		0.0010	-1.4656***	
	(0.0018)	(0.1271)		(0.0018)	(0.1271)	
Constant	7.5038	1.4479	3.4214***	7.5558	5.3649	3.3318***
	(4.7918)	(10.7882)	(0.2814)	(4.7876)	(11.4246)	(0.2813)
Observations	25,120	25,120	25,120	25,120	25,120	25,120
Adjusted R^2	0.4255	0.4564	0.6045	0.4290	0.4568	0.6071

Notes: Robust standard errors clustered at the firm level are reported in parentheses. Firm and year fixed effects are included. Price-to-Book Ratio is excluded from the Tobin's Q models. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.
Source: Author's elaboration

4.2 COVID Shock Intensity Indicator

Appendix Table 21 reports the mean values of the two alternative pandemic indicators for the China and EU samples, together with the results of the Welch t test, to examine their structural differences across samples. The results show that the mean value of COVID_Shock in the EU sample is 0.1403, significantly higher than 0.0002 in the Chinese sample. By contrast, the mean value of Stringency_Index in the Chinese sample is 55.2019, higher than 42.3894 in the EU sample. Both differences are significant at the 1% level. This result indicates that pandemic control patterns differ structurally between China and the EU. The EU sample is characterized by higher infection intensity and relatively lower policy stringency, while the Chinese sample is characterized by lower infection numbers and higher policy stringency.

Appendix Table 30: Mean Difference Tests for Pandemic Variables

Variable	EU Mean	China Mean	Difference EU - China	t-statistic
COVID_Shock	0.1403	0.0002	0.1401***	45.6539 (0.000)
Stringency_Index	42.3894	55.2019	-12.8126***	-30.8930 (0.000)

Notes: Difference is calculated as EU mean minus China mean. The values in parentheses are p values from Welch t tests. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.
Source: Author's elaboration.

This study uses COVIDShock_during as a pandemic shock indicator based on confirmed cases as a share of population, and further examines whether pandemic transmission intensity changes the relationship between digital transformation and firm performance. The results in Appendix Table 22 show that lnTextDigital × COVIDShock_during and ROD2 × COVIDShock_during are insignificant across all three performance indicators. This indicates that pandemic shock intensity measured by confirmed cases does not consistently moderate the relationship between the digitalization variables and firm performance.

This result may be related to structural differences in pandemic control patterns between China and the EU. Confirmed case numbers mainly reflect pandemic transmission intensity, but they may not fully represent the actual operational constraints faced by firms. Taking the Chinese sample as an example, infection numbers were relatively low in some periods, but government prevention and lockdown measures remained relatively strict, and firms could still face pressure from restrictions on labor mobility, supply chain disruptions, and offline operations. Although the EU sample had higher infection numbers, the way policies were implemented and the nature of firms' operational constraints differed from those in China. These features make it difficult for confirmed case numbers to consistently capture the actual operating pressure faced by firms in the combined China and EU sample. By contrast, the Stringency Index, as a direct measure of policy constraint intensity, has stronger proxy validity in the context of this study. Therefore, the regression results for COVIDShock_during serve as supplementary evidence from the perspective of pandemic transmission intensity and provide a comparison with the pandemic stage classification and policy stringency analysis in the main text.

Appendix Table 31: COVID Shock Intensity Results

Variables	ROA	ROE	Tobin's Q	ROA	ROE	Tobin's Q
	Panel A: lnTextDigital × COVID shock intensity			Panel B: ROD2 × COVID shock intensity		
lnTextDigital	0.3409***	2.2502**	-0.0184			
	(0.0939)	(0.9839)	(0.0180)			
ROD2				2.3389*	0.2093	0.3702***
				(1.2628)	(5.1818)	(0.0549)
COVID Shock during	-1.2501	75.3804	0.4907**	1.6429	19.9372	0.5734***
	(6.2514)	(48.7024)	(0.2229)	(2.5184)	(25.4426)	(0.1260)
lnTextDigital × COVID Shock	1.3917	-15.9020	0.0095			
	(1.6122)	(13.0655)	(0.0532)			
ROD2 × COVID Shock				2.7089	1.0384	-0.0108
				(2.6401)	(18.0303)	(0.1172)
Size	0.9049	4.0531**	-0.1929***	0.9047	4.6099***	-0.2126***

	(0.6348)	(1.6847)	(0.0378)	(0.6322)	(1.6013)	(0.0378)
Leverage	-21.0470***	-58.0967***	0.0419	-21.0172***	-58.2455***	0.0553
	(1.3879)	(10.7804)	(0.1533)	(1.3995)	(10.8350)	(0.1525)
Cash Flow Ratio	1.1813	1.3757	-0.0007	1.1444	1.3438	-0.0064
	(0.9057)	(1.3498)	(0.0238)	(0.8821)	(1.3407)	(0.0218)
Fixed Assets Ratio	-11.2048***	-8.3308	0.2799	-11.4894***	-9.7026	0.2816
	(1.0502)	(18.3567)	(0.2889)	(1.0491)	(17.8178)	(0.2873)
Price-to-Book Ratio	0.0012	-1.4652***		0.0010	-1.4655***	
	(0.0018)	(0.1271)		(0.0018)	(0.1272)	
Constant	7.3502	-0.9753	3.3821***	7.0612	2.0348	3.2483***
	(4.7512)	(11.3698)	(0.2809)	(4.7562)	(11.9357)	(0.2820)
Observations	25,120	25,120	25,120	25,120	25,120	25,120
Adj. R^2	0.4257	0.4567	0.6043	0.4285	0.4561	0.6070

Notes: Robust standard errors clustered at the firm level are reported in parentheses. Firm and year fixed effects are included. Price-to-Book Ratio is excluded from the Tobin's Q models. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.
Source: Author's elaboration