Estimating the Real Effect of Endogenous Digital Technology Adoption *

Mengyuan Cai Guiying Laura Wu

This version: March 2025
(click here for the latest version)

Abstract

This paper examines firms' optimal decisions on digital technology adoption. Using the endogenous treatment effect framework, we simultaneously address which firm characteristics are more likely to lead to adoption and what is the outcome of the adoption. We then apply the model to a panel of Chinese publicly listed firms. While observable firm characteristics suggest a positive selection consistent with existing literature, we find unobservable factors suggest a negative selection and lead to a "refugee sorting". Adopters would have underperformed in the non-adopting scenario but outperform after adoption. They gain market share, reduce production costs, enhance productivity, substitute capital for labour, and achieve a significantly higher profit. It is the positive correlation between non-adopting profit and adoption cost that hinders the digital adoption of those initially advantageous firms.

JEL Classification: O33, D22, L25, M15

Key Words: Digital Technology Adoption, Endogenous Treatment Effect Model, Self-Selection

*We would like to thank valuable comments from discussants and audiences at the Singapore Initiative on Digital Economics (SIDE) Workshop and the Economics Brownbag Seminar at Nanyang Technological University, and thank Qu Feng, Jie Gong, Xinyu Hua, Bo Li, Andrew Rhodes, Tat-How Teh, Jack Tong, Wenjie Wang, and Wu Zhu for their valuable comments. This research is supported by the Ministry of Education, Singapore, under its Academic Research Fund Tier 1 (RG54/22). Cai and Wu are from Nanyang Technological University, Singapore. Emails: mengyuan002@e.ntu.edu.sg (M. Cai), guiying.wu@ntu.edu.sg (G. Wu).

1 Introduction

The transformative impact of digital technologies, such as artificial intelligence (AI), big data, and machine learning, on modern life and production has spurred a burgeoning literature exploring their economic implications. While the stock market has responded quickly and enthusiastically on these technologies, the mechanisms driving their adoption and the consequences for firm performance remain underexplored. This paper addresses these gaps by examining the real effects of digital technology at the firm level, with a particular focus on the endogenous nature of adoption.

Existing studies provide valuable insights but leave critical questions unresolved. For instance, Acemoglu et al. (2023) highlights the distinction between causal effect and selection bias in advanced technology adoption, documenting that adoption tends to be concentrated in larger firms, which are often already more productive and growing faster. Their findings emphasize that selection effects dominate, raising concerns about the real adoption benefits. Other studies like Chen and Srinivasan (2023) reveals that firms engaging in digital activities enjoy substantial market value premiums but exhibit limited improvements in fundamental performance. These mixed findings underscore the importance of carefully disentangling the drivers and consequences of digital adoption, particularly for firms in emerging economies where adoption dynamics may diverge from those observed in advanced economies.

Our paper contributes to this growing body of literature by modeling digital technology adoption as an endogenous decision determined by firm characteristics and by simultaneously estimating its causal effects on firm performance. We first propose a general theoretical framework to understand firms' digital adoption choices, addressing the endogenous self-selection issue inspired by classic models like Borjas (1987). The framework models firms' costs and benefits with and without digital technology. A decision rule is derived from this general model, suggesting that firms adopt digital technology if and only if the benefits, in terms of enhanced concentration, demand expansion, and cost savings, exceed the digital adoption incurred uncertainty, fixed cost of adoption and the unit cost of subscription. To account for heterogeneity in adoption behavior across firms, we also allow

for potential correlation between non-adopting profit, adopting profit, and cost of adoption. This provides a foundation for empirical examination of the endogenous digital adoption patterns and their economic outcomes.

This paper addresses the challenge of measuring firm-level digital adoption by leveraging a textual analysis approach, focusing on annual reports of publicly listed firms in the Chinese stock market from 2009 to 2022. The analysis begins with constructing a dictionary of digital-related keywords, curated from official documents, policy reports, and word embedding techniques using the Bidirectional Encoder Representations from Transformers (BERT) model. These keywords are used to identify the occurrence and scope of digital adoption in firms' annual reports. Recognizing the limitations of raw keyword counts, we manually browse six thousands of annual reports sentences containing the digital keywords and propose a refinement algorithm to distinguish substantive actions from "cheap talk". The final measure, a binary indicator of digital adoption, classifies firms based on whether they report tangible digital engagement, which is less subject to measurement error. We finally observe that the proportion of digital technology adopter is around 15%. Comprehensive validity checks are conducted, which include keyword trends, industry-level adoption patterns, and firm-level evidence on positive associations with digital technology related patents, software assets, R&D intensity, and amortizations of intangible assets. These findings confirm that our text-based measure to a large extent reflects the digital technology engagement in firms, contributing to the growing literature on the measure of digital transformation.

To empirically estimate the effects of digital adoption on firm performance, we develop an endogenous treatment effect framework that accounts for firm's self-selection into digital adoption. Our general model specifies the adoption decision through a binary choice equation and allows for heterogeneous outcome responses, where firm profit depends on both observable characteristics and unobservable components that may correlate with the adoption decision. We then propose a simplified version that imposes uniform estimates of characteristics across adoption status while maintaining endogenous treatment effects, offering computational efficiency in estimating average treatment effect (ATE) and treatment effect on the treated (ATET). For identification, we employ

a control function approach and leverage three instrumental variables: the length of firms' annual reports, city-level broadband internet penetration rates, and CEO's IT-related background. These instruments influence firms' digital adoption decisions but are plausibly exogenous to post-adoption performance after controlling for firm characteristics and rich fixed effects (FEs).

Our empirical analysis yields several key findings. The Probit model reveals that digital adopters are typically larger firms with higher cash flow, lower leverage, reduced sales volatility, and more skilled labour, while also showing that our instrumental variables are relevant predictors of adoption. Examining the performance outcomes using both maximum likelihood estimation (MLE) and control function (CF) approaches, we find significant positive treatment effects of digital adoption on firm profits, accompanied by interesting patterns of self-selection. Specifically, we document a positive ATE on logarithm of profit to be 0.255 and a larger treatment effect on the treated of 0.670, where the predicted non-adopting logarithm of profit is 19.768. The increase in profit is primarily driven by demand expansion, cost reduction, and productivity enhancement. Moreover, this technological shift also leads to capital substituting for labour, indicating that digital technologies transform production processes by altering firms' optimal input combination. Additionally, we observe short-term boosts in market valuation, as evidenced by a sharp rise in Tobin's Q during the adoption year.

When we decompose the group mean difference (GMD) in profit into ATET and selection bias (SB), we find the observable firm characteristics suggest a positive selection. This is consistent with existing literature on robots and automation using firm-level data from advanced economies, such as Koch et al. (2021); Acemoglu and Restrepo (2020); Acemoglu et al. (2023); Anderton et al. (2020). However, the unobservable factors indicate a negative selection in our sample. Such factors dominate the observable factors and lead to the "refugee sorting" pattern, where adopters would have underperformed in the non-adopting scenario but outperform after adoption. This suggests that digital technology serves as a transformative opportunity for firms with initially disadvantaged positions. The positive correlation between non-adopting profit and cost of adoption also explains why digital technology is not widely adopted in our sample.

The rest of paper is organized as follows. Section 2 proposes a theoretical framework of digital adoption choice that accounts for the endogenous self-selection issue. Section 3 presents the empirical specification for the endogenous treatment effect model. Section 4 describes data employed in this paper. Section 5 shows the estimation results. Section 6 investigates the mechanisms. Section 7 concludes the paper.

2 A Theoretical Framework

2.1 Basic Setup

Our theoretical model shares the same conceptual framework as Hjort and Tian (2024) where we allow the digital technology to affect firms' optimal decisions through both demand and supply forces in a general way. We begin by considering a firm's optimal pricing under industry equilibrium, following Bresnahan (1989) and Weyl and Fabinger (2013). The firm's optimal price, P^* , is given by:

$$P^* = \frac{e}{e-\theta}m, \tag{1}$$

where $e > 1$ is the price elasticity of demand faced by an individual firm, and $\theta \in [0, 1]$ is the conduct parameter, capturing the degree of market competition. A value of $\theta = 1$ corresponds to a monopoly, while $\theta = 0$ is perfect competition. Lastly, m is the marginal cost of production.

In this setup, the firm's markup is given by $\frac{\theta}{e-\theta} = \frac{1}{\frac{e}{\theta}-1}$. The firm's profitability is given by $\frac{PQ-mQ}{PQ} = 1 - \frac{m}{p} = 1 - \frac{e-\theta}{\theta} = \frac{\theta}{e}$. This shows that a firm in industry with a higher value of θ has greater market power to set a higher markup and gets the higher profitability.

Given the demand function $Q = dP^{-e}$, where d is a demand shifter, firm's maximized profit is:

$$\begin{aligned} \Pi &= P^* \cdot Q^* - m \cdot Q^* \\ &= (P^* - m)Q^* \\ &= (\frac{e}{e-\theta} - 1)md(\frac{e}{e-\theta}m)^{-e} \\ &= am^{1-e}d, \end{aligned} \tag{2}$$

where $a \equiv \theta(e-\theta)^{e-1}e^{-e}$. Here $\frac{\partial \Pi}{\partial a} > 0$ and $\frac{\partial a}{\partial \theta} = e^{1-e}(e-\theta)^{e-2}(1-\theta) > 0$.

It is transformed in log-linear form as:

$$\ln \Pi = \ln a + (1-e)\ln m + \ln d. \tag{3}$$

Let A denote the adoption of digital technology in areas such as production, management, and marketing. Specifically,

$$A = \begin{cases} 0: \text{ No digital adoption} \\ 1: \text{ Digital adoption} \\ R^{+}: \text{Intensity of usage, more generally.} \end{cases}$$

The relationship between the advancement of digital technology A and price elasticity of demand e is debated in the literature. On the one hand, digital technologies like online marketing reduce consumers' search costs (Goldfarb and Tucker, 2019), fostering competition and potentially increasing the price elasticity of demand. For instance, Granados et al. (2012) show that online demand, measured from millions of airline ticket sales, is more price-elastic compared to offline channels. On the other hand, technologies such as digital advertising (Greenwood et al., 2024), big data, and recommendation algorithms driven by AI (Ban and Keskin, 2021; Dubé and Misra, 2023) can create more personalized and targeted product offerings. This, in turn, makes consumers more sticky to a particular product or brand, reducing the price elasticity faced by firms. Furthermore, as summarized by Bessen (2018), demand elasticity can vary at different stages of market development. AI-driven automation, for example, tends to reduce production costs, leading to lower prices and higher demand elasticity in the early stages. However, as prices continue to fall and markets become saturated, demand becomes less elastic, and further price reductions have a diminishing impact on demand.

Given the conflicting evidence, we assume that price elasticity of demand, e, remains constant with respect to digital technology adoption, A. While our prior assumptions strongly suggest that digital adoption increases market demand, i.e., $\frac{\partial d(A)}{\partial A} > 0$, and reduces production costs, i.e., $\frac{\partial m(A)}{\partial A} < 0$, we leave the impact of digital adoption on industry competition, θ, open to interpretation.

Digital adoption incurs both a fixed cost F and a per period subscription cost $s(A)$, where $\frac{\partial s(A)}{\partial A} > 0$. The fixed cost F can be a lump sum fee of constructing a data center, and can also be an adjustment cost of digital adoption, e.g., the cost of labour union resistance to machines replacing workers. We assume, for modeling convenience, that the subscription fee is proportional to profit.

If firms choose no digital adoption, the firm value is:

$$V^0(A=0) = \frac{\Pi(0)}{r_0} = \frac{1}{r_0} a(0) m(0)^{1-e} d(0). \tag{4}$$

where V^0 is the present value of non-adopting firms considering all future cash flow $\Pi(0)$ with a discount rate r_0, according to the perpetuity formula.

In contrast, if firms choose to adopt digital technology, the firm value is:

$$V^1\left(A>0\right) = -F + \frac{[1-s(A)]\Pi(A)}{r_0(1+\delta)} = -F + \frac{[1-s(A)]}{r_0(1+\delta)} a(A) m(A)^{1-e} d(A). \tag{5}$$

where V^1 is the present value of adopting firms considering the upfront paid adoption cost F, and future cash flow $\Pi(A)$ after deducting the per period subscription fee $s(A)$, with a discount rate $r_0(1+\delta)$ where δ is the risk premium incurred by digital adoption.

In this setup, firms choose digital adoption if and only if:

$$V^1(A>0) \geqslant V^0(A=0). \tag{6}$$

And given adopting, the optimal intensity of usage is given by:

$$-\frac{\partial s(A)}{\partial A} + \frac{\partial \ln a(A)}{\partial A} - (e-1)\frac{\partial \ln m(A)}{\partial A} + \frac{\partial \ln d(A)}{\partial A} = 0. \tag{7}$$

For simplicity while maintaining its generality, we do not distinguish firms with $A = R^+$ and $A = 1$ and consider the values of two types of firms with $A = 0$ and $A = 1$:

$$V^0 = \frac{\Pi(0)}{r_0}, \tag{8}$$

$$V^1 = -F + \frac{[1-s(1)]\Pi(1)}{r_0(1+\delta)}. \tag{9}$$

Firms choose to go digital if and only if $V^1 - V^0 \geqslant 0$, i.e.,

$$\frac{[1-s(1)]\Pi(1)}{r_0(1+\delta)} - \frac{\Pi(0)}{r_0} \geqslant F. \tag{10}$$

Equivalently,

$$[1-s(1)]\Pi(1) - (1+\delta)\Pi(0) \geqslant r_0(1+\delta)F = (1+\delta)f \cdot \Pi(0), \tag{11}$$

where $f \equiv \frac{r_0 F}{\Pi(0)}$, which transforms the fixed cost of digital adoption as an annual flow of the opportunity cost of risk-free capital as a proportion of non-adopting profit.

Thus, we can show that the digital adoption rule is given by:

$$\begin{aligned}
[1-s(1)]\Pi(1) &\geqslant (1+\delta)(1+f)\Pi(0) \\
[1-s(1)]a(1)d(1)m(1)^{1-e} &\geqslant (1+\delta)(1+f)a(0)d(0)m(0)^{1-e} \\
(1-s)a_1 d_1 m_1^{1-e} &\geqslant (1+\delta)(1+f)a_0 d_0 m_0^{1-e} \\
\ln(1-s) + \ln a_1 + \ln d_1 + (1-e)\ln m_1 &\geqslant \ln(1+\delta) + \ln(1+f) + \ln a_0 + \ln d_0 + (1-e)\ln m_0 \\
-s + \ln a_1 + \ln d_1 - (e-1)\ln m_1 &\geqslant \delta + f + \ln a_1 + \ln d_0 - (e-1)\ln m_0 \\
\ln\frac{a_1}{a_0} + \ln\frac{d_1}{d_0} + (e-1)\ln\frac{m_0}{m_1} &\geqslant \delta + f + s \\
\ln\Pi_1 - \ln\Pi_0 &\geqslant \delta + f + s.
\end{aligned} \tag{12}$$

The intuition behind this rule is straightforward. Firms choose digitization if and only if the benefits from industry concentration, market demand, and the cost reduction exceed the digital adoption incurred uncertainty, fixed cost of adoption and the unit cost of subscription. Moreover, note that the incremental profit is increasing in the demand elasticity e. It implies that the more elastic the demand faced by firms, the more willing to adopt for firms.

For simplicity of notations, we denote $\pi_{i1} \equiv \ln\Pi_{i1}$, $\pi_{i0} \equiv \ln\Pi_{i0}$, and $c_i \equiv \delta_i + f_i + s_i$, and then the digital adoption rule becomes:

$$A_i = \mathbf{1}\{\pi_{i1} - \pi_{i0} \geq c_i\}, \tag{13}$$

where **1** is an indicator function equals 1 if the condition holds, otherwise 0.

2.2 Probability of Digital Adoption

To characterize the endogenous nature of digital technology adoption, we follow Borjas (1987, 1991), a classic model on international migration decision, by allowing potential correlation between the benefits and costs of the adoption. Specifically, we model profit and cost as the sum of a population mean and an idiosyncratic component. The population mean represents the observable portion explained by firm characteristics, while the idiosyncratic error captures unobservable factors. Specifically, if firm i does not adopt digital technology, the natural logarithm of its profit is given by:

$$\pi_{i0} = \mu_0 + \varepsilon_{i0}. \tag{14}$$

In contrast, if firm i adopts digital technology, the natural logarithm of its profit is expressed as:

$$\pi_{i1} = \mu_1 + \varepsilon_{i1}. \tag{15}$$

The cost of digital adoption is written as:

$$c_i = \mu_c + \varepsilon_{ic}. \tag{16}$$

In this framework, μ_0, μ_1, and μ_c represent the potential means of the natural logarithm of profit for non-adopting firms, adopting firms, and the cost of digital adoption, respectively. These means are generally functions of observable firm characteristics such as firm size, cash flow, industry, and growth prospects, which will be explored further in the estimation sections.

The terms ε_{i0}, ε_{i1}, and ε_{ic} capture firm-specific factors that are unobservable to researchers. We assume these unobservable factors follow a trivariate normal distribution as follows:

$$\begin{pmatrix} \varepsilon_{i0} \\ \varepsilon_{i1} \\ \varepsilon_{ic} \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_0^2 & \rho_{01}\sigma_0\sigma_1 & \rho_{0c}\sigma_0\sigma_c \\ \rho_{01}\sigma_0\sigma_1 & \sigma_1^2 & \rho_{1c}\sigma_1\sigma_c \\ \rho_{0c}\sigma_0\sigma_c & \rho_{1c}\sigma_1\sigma_c & \sigma_c^2 \end{pmatrix} \right], \tag{17}$$

where in the variance-covariance matrix, σ_0, σ_1, and σ_c represent the standard deviations of ε_{i0}, ε_{i1}, and ε_{ic} in the population. The terms ρ_{01}, ρ_{0c}, and ρ_{1c} denote the pairwise correlation coefficients between ε_{i0}, ε_{i1}, and ε_{ic}.

Under the normality assumption and following the digital adoption rule in Equation (13), the probability of digital adoption is expressed as:

$$\begin{aligned} P &= \Pr\left[\mu_1 + \varepsilon_{i1} - \mu_0 - \varepsilon_{i0} \geqslant \mu_c + \varepsilon_{ic}\right] \\ &= \Pr\left[\varepsilon_{i1} - \varepsilon_{i0} - \varepsilon_{ic} \geqslant -\left(\mu_1 - \mu_0 - \mu_c\right)\right] \\ &= \Pr\left[v_i/\sigma_v \geqslant -\left(\mu_1 - \mu_0 - \mu_c\right)/\sigma_v\right] \\ &= 1 - \Phi(\omega) \\ &= \Phi(-\omega) \end{aligned} \tag{18}$$

where

$$v_i = \varepsilon_{i1} - \varepsilon_{i0} - \varepsilon_{ic}, \tag{19}$$

and

$$-\omega = \left(\mu_1 - \mu_0 - \mu_c\right)/\sigma_v. \tag{20}$$

Here, $v_i \sim N(0, \sigma_v^2)$, and Φ is the standard normal cumulative distribution function.

According to Equation (18), the probability of digital adoption is affected by three main factors. First, it is positively related to the expected adopting profit, μ_1, i.e., $\partial P/\partial \mu_1 > 0$. Second, it is negatively related to the expected non-adopting profit, μ_0, i.e., $\partial P/\partial \mu_0 < 0$. Third, it is negatively related to the expected cost of adoption, μ_c, i.e., $\partial P/\partial \mu_c < 0$.

By modelling μ_0, μ_1, and μ_c as functions of observable firm characteristics $\boldsymbol{X}$, the effects of these characteristics on the probability of digital adoption can be derived as:

$$\frac{\partial P}{\partial \boldsymbol{X}} = \frac{d\Phi(-\omega)}{d(-\omega)} \cdot \frac{\partial(-\omega)}{\partial \boldsymbol{X}}, \tag{21}$$

where $\frac{d\Phi(-\omega)}{d(-\omega)} > 0$. Therefore, the sign of $\frac{\partial P}{\partial \boldsymbol{X}}$ depends on the sign of $\frac{\partial(-\omega)}{\partial \boldsymbol{X}}$.

Specifically, we have:

$$\frac{\partial(-\omega)}{\partial \boldsymbol{X}} = \frac{\partial[\mu_1(\boldsymbol{X}) - \mu_0(\boldsymbol{X}) - \mu_c(\boldsymbol{X})]}{\partial \boldsymbol{X}} \cdot \frac{1}{\sigma_v}.$$

2.3 Self-Selection

The decision to adopt digital technology is endogenous, influenced by firm-specific factors. This is mathematically represented by $v_i = \varepsilon_{i1} - \varepsilon_{i0} - \varepsilon_{ic}$, where v_i affects the adoption decision as in Equation (18). At the same time, ε_{i0} and ε_{i1} affect π_{i0} and π_{i1}, which are the potential profits in the absence and presence of digital adoption, respectively.

To address this endogeneity issue, we introduce the following notations. Let $E[\pi_{i1}]$ represent the population mean of adopting profit, and $E[\pi_{i0}]$ represent the population mean of non-adopting profit. By definition, $E[\pi_{i1}] = \mu_1$ and $E[\pi_{i0}] = \mu_0$. We denote the mean profit of adopting firms as $E[\pi_{i1}|A_i = 1]$, and the *de jure* counterfactual mean profit of adopting firms if they had not adopted digital technology as $E[\pi_{i0}|A_i = 1]$. Note that the *de facto* adopting firms are featured by $A_i = 1$ and the non-adopting firms are featured by $A_i = 0$.

Assuming normality, the conditional means for these two subgroups can be written as:

$$E\left[\pi_{i0} \mid A_i = 1\right] = \mu_0 + \frac{\sigma_0\sigma_1}{\sigma_v}\left[\left(\rho_{01} - \frac{\sigma_0}{\sigma_1}\right) - \rho_{0c}\frac{\sigma_c}{\sigma_1}\right] \cdot h, \tag{22}$$

and

$$E\left[\pi_{i1} \mid A_i = 1\right] = \mu_1 + \frac{\sigma_0\sigma_1}{\sigma_v}\left[\left(\frac{\sigma_1}{\sigma_0} - \rho_{01}\right) - \rho_{1c}\frac{\sigma_c}{\sigma_0}\right] \cdot h, \tag{23}$$

where $h = \phi(-\omega)/\Phi(-\omega) = \phi(\omega)/P$, with ϕ denoting the probability density function of the standard normal distribution. The term h is commonly referred to as the hazard rate, or inverse Mills ratio, as discussed in Heckman (1979).

For simplicity of notation, following Borjas (1991), we define:

$$S_0 \equiv \frac{\sigma_0\sigma_1}{\sigma_v}[(\rho_{01} - \frac{\sigma_0}{\sigma_1}) - \rho_{0c}\frac{\sigma_c}{\sigma_1}]h, \tag{24}$$

as the difference between the population mean of non-adopting profit, μ_0, and the counterfactual profit of adopting firms if they had not adopted, $E[\pi_{i0}|A_i = 1]$.

Similarly, we define:

$$S_1 \equiv \frac{\sigma_0\sigma_1}{\sigma_v}[(\frac{\sigma_1}{\sigma_0} - \rho_{01}) - \rho_{1c}\frac{\sigma_c}{\sigma_0}]h, \tag{25}$$

as the difference between the population mean of adopting profit, μ_1, and the *de facto* profit of adopting firms, $E[\pi_{i1}|A_i = 1]$.

Intuitively, the signs of S_0 and S_1 depict the firms' positions in profit distribution of non-adopting and adopting circumstances. In detail, $S_0 > 0$ indicates that if no firm were to adoptdigital technology, the potentially adopting firms ($A_i = 1$) would, on average, be positioned to the right of the profit distribution, compared with the *de facto* non-adopting firms ($A_i = 0$). Conversely, $S_0 < 0$ would mean the opposite. Similarly, $S_1 > 0$ means that, were all firms to adopt digital technology, the *de facto* adopting firms ($A_i = 1$) would, on average, be positioned to the right of the profit distribution, compared with the potentially non-adopting firms ($A_i = 0$). And $S_1 < 0$ would mean the opposite.

In Borjas (1991), self-selection is featured by the signs of S_0 and S_1. Positive selection occurs if $S_0 > 0$ and $S_1 > 0$. This case requires the *de facto* adopting firms are on average more profitable in both adopting and non-adopting circumstances. Negative selection refers to $S_0 < 0$ and $S_1 < 0$. It depicts that the *de facto* adopting firms are on average less profitable in both adopting and non-adopting circumstances. Lastly, the case of $S_0 < 0$ and $S_1 > 0$ is called "refugee sorting", as it indicates that the *de facto* adopting firms are on average less profitable in the non-adopting circumstances but more profitable in the adopting circumstances.

2.4 The Treatment Effect of Digital Adoption

Suppose firms are continuous and there exists a marginal firm m that is indifferent to adopting digital technology, given its firm characteristics. For this marginal firm m, the digital adoption rule in Equation (13) implies the following equilibrium condition:

$$E[\pi_{m1}] - E[\pi_{m0}] = c_m. \tag{26}$$

Equation (26) represents the market equilibrium condition. If this condition is violated, for example, when an additional firm decides to adopt digital technology, say, AI, the increased demand for

AI specialists will raise the cost of adoption. This increase in cost, represented by a larger c_m , discourages further adoption of digital technology by other firms. The adjustment process continues until the marginal firm is once again indifferent to adopting digital technology.

We denote the average treatment effect and the average treatment effect on the treated by ATE and $ATET$, respectively. At the point of deciding whether to adopt, the marginal firm makes its best estimates: $E[\pi_{m0}] = \mu_0$ and $E[\pi_{m1}] = \mu_1$. Therefore, using Equations (14), (15), and (26), the ATE of digital adoption can be expressed as:

$$ATE = E[\pi_{i1}] - E[\pi_{i0}] = \mu_1 - \mu_0 = c_m. \tag{27}$$

Additionally, based on Equations (22) and (23), the $ATET$ of digital adoption is given by:

$$ATET = E[\pi_{i1} - \pi_{i0}|A_i = 1] = (\mu_1 - \mu_0) + (S_1 - S_0) = c_m + (S_1 - S_0)\,. \tag{28}$$

Thus, an important implication of our theoretical framework is that the ATE reflects the cost of adoption for the marginal firm, c_m, while the $ATET$ captures both the cost of adoption and the relative positioning of adopting firms in the profit distributions of adopters and non-adopters, represented by $(S_1 - S_0)$.

We also denote selection bias as SB, which is defined as the difference in the expected non-adopting profits between the firms that are *de facto* adopters ($A_i = 1$) and non-adopters ($A_i = 0$):

$$\begin{aligned} SB &= E[\pi_{i0}|A_i = 1] - E[\pi_{i0}|A_i = 0] \\ &= (\mu_0 + S_0) - \left(\mu_0 - \frac{P}{1-P}S_0\right) \\ &= \frac{S_0}{1-P}. \end{aligned} \tag{29}$$

According to Equation (29), SB shares the same sign as S_0. Consequently, SB provides insights into the direction of selection prior to digital adoption. It is important to note that only the group mean difference denoted by GMD between the adopters ($A_i = 1$) and the non-adopters ($A_i = 0$) groups can be observed, which is defined as:

$$
\begin{aligned}
GMD &= E[\pi_{i1}|A_i = 1] - E[\pi_{i0}|A_i = 0] \\
&= (\mu_1 + S_1) - \left(\mu_0 - \frac{P}{1-P} S_0\right) \\
&= c_m + (S_1 - S_0) + \left(\frac{S_0}{1-P}\right) \\
&= ATET + SB
\end{aligned} \tag{30}
$$

Equation (30) shows that GMD can be decomposed into the sum of $ATET$ and SB. As highlighted by Acemoglu et al. (2023), the observed profit premium ($GMD > 0$) between automation-adopting and non-adopting firms in the U.S. could stem from two factors: either a positive causal (treatment) effect on the adopting firms ($ATET > 0$), or a positive selection bias ($SB > 0$) before adopting automation. If the selection bias is negative, the actual treatment effect would be even larger than the observed profit premium. To uncover the real effect of digital adoption, an endogenous treatment effect model is proposed in this paper, empirically decomposing the observed GMD into $ATET$ and SB.

3 An Endogenous Treatment Effect Model

3.1 A General Model

In this section, we develop an empirical specification based on our theoretical framework. First of all, a general model is proposed to describe the endogenous decision to digital technology adoption and the subsequent performance outcomes. The empirical specification for the endogenous treatment effect model is featured as:

$$A_{it} = \mathbf{1}\{\boldsymbol{X_{i,-t}}'\boldsymbol{\alpha_1} + \boldsymbol{Z_{i,-t}}'\boldsymbol{\alpha_2} + v_{i,-t} > 0\}, \tag{31}$$

$$y_{i\bar{t}} = A_{it}y_{1i\bar{t}} + (1 - A_{it})y_{0i\bar{t}}, \tag{32}$$

$$y_{0i\bar{t}} = \boldsymbol{X_{i\bar{t}}}'\boldsymbol{\beta_{10}} + \varepsilon_{0i\bar{t}}, \tag{33}$$

$$y_{1i\bar{t}} = \boldsymbol{X_{i\bar{t}}}'\boldsymbol{\beta_{11}} + \varepsilon_{1i\bar{t}}, \tag{34}$$

$$cov[\varepsilon_{ji\bar{t}}, v_{i,-t}] \neq 0 \text{ for } j \in \{0, 1\}. \tag{35}$$

Note that the subscript t denotes adoption batch year t; $-t$ represents years before the year t, and $\bar{t}$ is years after year t [1].

Equation (31) formalizes the digital adoption rule from Equations (13) and (18). Here, A_{it} is a binary treatment indicator that equals 1 if firm i belongs to the treatment group of batch t, and 0 otherwise. The adoption probability at year t depends on firm characteristics ($\boldsymbol{X}_{i,-t}$) and instrumental variables ($\boldsymbol{Z}_{i,-t}$) from previous years, reflecting that adoption decisions typically rely on information available 2-3 years prior. Equation (32) specifies $y_{i\bar{t}}$ as the observed post-adoption firm performance taking averages over years, while $y_{1i\bar{t}}$ and $y_{0i\bar{t}}$ represent potential outcomes under treatment and control conditions, respectively.

According to Equations (33) and (34), each potential outcome is determined by a vector of observed firm-specific characteristics, $\boldsymbol{X_{i\bar{t}}}$, along with an unobserved random component, $\varepsilon_{ji\bar{t}}$, $j \in \{0, 1\}$, that encompasses factors affecting firm performance but unobservable to researchers. Similarly, the adoption decision in Equation (31) is determined by a set of regressors, $\boldsymbol{X_{i,-t}}$ and $\boldsymbol{Z_{i,-t}}$, and an unobserved component, $v_{i,-t}$. Here, $\boldsymbol{Z_{i,-t}}$ includes instrumental variables that affect the decision to adopt digital technology but do not directly affect firm performance after digital adoption. We assume that $E[v_{i,-t}|\boldsymbol{X_{i,-t}}, \boldsymbol{Z_{i,-t}}] = 0$, enabling consistent estimation of α_1 and α_2 in the Probit model outlined in Equation (31).

[1] In Section 4.2, we detail that the treatment group for batch t is defined over a 5-year horizon. Treatment firms have no digital adoption in years $t-2$ and $t-1$ but adopt in years t, $t+1$, and $t+2$. The control group consists of firms with no digital adoption throughout years $t-2$ to $t+2$.

Equation (35) addresses the endogeneity issue by acknowledging that unobserved components $\varepsilon_{ji\bar{t}}$, $j \in \{0,1\}$ in Equations (33) and (34) are correlated with the treatment status through $v_{i,-t}$, where $v_i = \varepsilon_{i1} - \varepsilon_{i0} - \varepsilon_{ic}$ as in Equation (19). For example, firms with superior AI expertise are more likely to both adopt digital technology ($v_i > 0$) and achieve better performance ($\varepsilon_{i1} > 0$) post-adoption.

3.2 Link to the Theory

Putting together, to align the empirical model with our theoretical framework, we have the following specification:

$$\begin{aligned} \mu_0 &= E[\boldsymbol{X_i}'\boldsymbol{\beta_0}], \\ \mu_1 &= E[\boldsymbol{X_i}'\boldsymbol{\beta_1}], \\ -w &= (\mu_1 - \mu_0 - \mu_c)/\sigma_v = E[(\boldsymbol{X_i}'\boldsymbol{\alpha_1} + \boldsymbol{Z_i}'\boldsymbol{\alpha_2})/\sigma_v], \\ v_i &= \varepsilon_{i1} - \varepsilon_{i0} - \varepsilon_{ic}. \end{aligned}$$

We assume that the vector of error terms $(\varepsilon_{i0}, \varepsilon_{i1}, v_i)'$ follows a trivariate normal distribution:

$$\begin{pmatrix} \varepsilon_{i0} \\ \varepsilon_{i1} \\ v_i \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_0'^2 & \rho_{01}'\sigma_0'\sigma_1' & \rho_{0v}'\sigma_0' \\ \rho_{01}'\sigma_0'\sigma_1' & \sigma_1'^2 & \rho_{1v}'\sigma_1' \\ \rho_{0v}'\sigma_0' & \rho_{1v}'\sigma_1' & 1 \end{pmatrix} \right]. \tag{36}$$

Here, σ_0' and σ_1' represent the dispersion of ε_{i0} and ε_{i1} for the subgroup of control ($A_i = 0$) and treated ($A_i = 1$), respectively. The term σ_v' is normalized to 1, as it is neither identifiable nor relevant for inference. The parameter ρ_{01}' is not identified since it is impossible to observe a firm both adopting and not adopting simultaneously.

A comparison of the two matrices (17) and (36) offers an important and valuable insight: despite the limitation that researchers can only observe firm performance in one scenario and cannot directly measure adoption cost, the endogenous treatment effect framework allows for the indirect estimation of the distribution of firm performance and adoption cost across the whole population by the potential outcomes predicted by the empirical model.

3.3 A Simple Model

We propose a simplified version of the general model that consolidates Equations (32), (33), and (34) into:

$$y_{i\bar{t}} = \boldsymbol{X_{i\bar{t}}}'\boldsymbol{\beta_1} + \gamma A_{it} + \varepsilon_i, \tag{37}$$

with the variance-covariance matrix of error terms simplified to:

$$\begin{bmatrix} \sigma'^2 & \rho'\sigma' \\ \rho'\sigma' & 1 \end{bmatrix} \tag{38}$$

This simple model imposes uniform contribution of characteristics $\boldsymbol{X_{i\bar{t}}}$ to profit across both scenarios, with the treatment indicator A_{it} serving as the only difference in the profit. This structure yields $ATE = E[y_{1i\bar{t}} - y_{0i\bar{t}}] = \gamma$. In addition, $E[\varepsilon_{1i} - \varepsilon_{0i} | A_i = 1] = 0$, or equivalently, $S_0 = S_1 = \rho\sigma h$ in Equations (24) and (25). Consequently, $ATET = ATE = \gamma$ in this simple model.

This simplified structure offers the key advantage of computational efficiency in estimating $ATET$ and ATE through a single parameter γ. Moreover, note that the identical values of $ATET$ and ATE stem not from the absence of selection bias, but from the model's assumption that selection bias operates with equal strength in both adoption and non-adoption scenarios. Despite its simplified nature, the model retains endogenous treatment effects, evidenced by non-zero off-diagonal elements in (38), necessitating appropriate endogeneity controls for consistent γ estimation.

3.4 Identification

While MLE or Heckit methods can provide consistent estimates $\boldsymbol{\beta_{10}}$ and $\boldsymbol{\beta_{11}}$ (or $\boldsymbol{\beta_1}$ and γ) in the general (simple) model, under trivariate (bivariate) normality assumptions on $(\varepsilon_{i0}, \varepsilon_{i1}, v_i)'$ $((\varepsilon_i, v_i)')$, we implement CF approach (Wooldridge, 2010) as a more flexible alternative. Using the simple model for illustration, the main idea of the CF approach is to model the correlation between endogenous treatment A_i and the error term ε_i by projecting ε_i on A_i, $\boldsymbol{X_i}$, and identifying variables

$\boldsymbol{Z_i}$. Given that $A_i = E[A_i|\boldsymbol{X_i}, \boldsymbol{Z_i}] + (A_i - E[A_i|\boldsymbol{X_i}, \boldsymbol{Z_i}])$, we write:

$$\begin{aligned} & E[\varepsilon_i|A_i, \boldsymbol{X_i}, \boldsymbol{Z_i}] \\ = \; & E[\varepsilon_i|E(A_i|\boldsymbol{X_i}, \boldsymbol{Z_i}) + (A_i - E(A_i|\boldsymbol{X_i}, \boldsymbol{Z_i})), \boldsymbol{X_i}, \boldsymbol{Z_i}] \\ = \; & E[\varepsilon_i|A_i - E(A_i|\boldsymbol{X_i}, \boldsymbol{Z_i})] = E[\varepsilon_i|v_i] = v_i\beta_2, \end{aligned}$$

where $v_i = A_i - E[A_i|\boldsymbol{X_i}, \boldsymbol{Z_i}]$.

Therefore, we have:

$$E[y_i|A_i, \boldsymbol{X_i}, \boldsymbol{Z_i}] = \boldsymbol{X_i'\beta_1} + \gamma A_i + v_i\beta_2. \tag{39}$$

For implementation, v_i is proxied by $\widehat{v}_i = A_i - \Phi(\boldsymbol{X_i'}\widehat{\boldsymbol{\alpha}_1} + \boldsymbol{Z_i'}\widehat{\boldsymbol{\alpha}_2})$ from the Probit regression (31), and the coefficient β_2 serves as a selection bias indicator.

In the general model, the CF approach extends to:

$$y_{ij} = \boldsymbol{X_i'\beta_{1j}} + v_i\beta_{2j} + e_{ij}, \; j \in \{0, 1\}, \tag{40}$$

with β_{20} and β_{21} indicating selection bias under non-adoption and adoption, respectively.

The treatment effects should be calculated as:

$$\begin{aligned} \widehat{ATE} &= \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{X_i'}\left(\hat{\boldsymbol{\beta}}_{\mathbf{11}} - \hat{\boldsymbol{\beta}}_{\mathbf{10}}\right), \\ \widehat{ATET} &= \frac{1}{\sum_{i=1}^{N} A_i}\sum_{i=1}^{N} \boldsymbol{X_i'}\left(\hat{\boldsymbol{\beta}}_{\mathbf{11}} - \hat{\boldsymbol{\beta}}_{\mathbf{10}}\right) A_i + \frac{1}{\sum_{i=1}^{N} A_i}\sum_{i=1}^{N} \widehat{v}_i\left(\hat{\beta}_{21} - \hat{\beta}_{20}\right) A_i. \end{aligned}$$

The above-mentioned parameters can be estimated either through one-step GMM or a two-step CF procedure combining Probit and OLS estimation.

3.5 Instrumental Variables

To implement our identification strategy, we use three instrumental variables: the length of the annual report, the broadband internet penetration rate in the city where the firm is located, and

whether the firm is led by an CEO with IT-related background. These variables affect the firm's decision to adopt digital technology as outlined in Equation (31) but are uncorrelated with its performance post-adoption in Equations (33) and (34).

First, length of the annual report is positively associated with the probability of a firm being classified as treated. This is because our treatment group is defined based on whether firms report substantial digital engagement in their annual reports. Longer reports are more likely to be identified by machine and to pass our refinement algorithm. Furthermore, the verbosity of the board secretary is unlikely to affect firm performance after digital adoption, especially when controlling for firm-specific characteristics and a rich set of FEs.

Second, city-level broadband internet penetration rate reflects the level of Information and Communication Technology (ICT) infrastructure. On the one hand, better ICT infrastructure like broadband internet, serving as the complementary goods to digital technologies, makes it easier and more cost-effective for local firms to adopt digital technologies such as cloud computing. On the other hand, higher internet penetration in the city introduces more digitized customers and suppliers, further incentivizing firms to adopt digital technology. Thus, firms located in areas with higher broadband penetration are more likely to embrace digital technologies. Additionally, given the inclusion of firm-specific characteristics and FEs, this city-level internet penetration is less likely to directly affect firm performance.

Third, we collect data on whether the CEO has an IT-related background. In literature, this information is usually gathered by manually reviewing CEO resumes and assessing their work experience and educational qualifications (Haislip and Richardson, 2018). According to the 2004 Guidelines on the Content and Format of Annual Reports for Publicly Listed Firms issued by the China Securities Regulatory Commission, listed firms are required to disclose the primary work experience of directors, supervisors, and senior executives in their annual reports. Therefore, we classify a CEO as having an IT-related background if their education or professional experience is related to information technology or information management. We expect firms led by CEOs with

IT-related backgrounds are on average more likely to adopt digital technology. In addition, such IT background is less likely to directly affect firm profit after digital adoption.

4 Data

4.1 Digital Technology Adoption

One of the key challenges to be addressed in this paper is how to measure the firm-level digital adoption. In the literature, digital adoption is mainly measured in three ways. First, the most direct and ideal source of data on adoption status comes from enterprise surveys and census datasets. For example, Acemoglu et al. (2022) utilize the 2019 Annual Business Survey, which includes a new module on the use of AI, robotics, dedicated equipment, specialized software, and cloud computing. Similarly, to investigate digital technology adoption such as e-commerce, remote work, and electronic information systems, Cong et al. (2024) combine administrative firm registration data with offline and online surveys of small and medium-sized enterprise (SME) owners in China.

A second method uses proxies that indicate a firm's digital engagement. Examples include the ratio of IT personnel (Babina et al., 2024), investments in ICT (Becchetti et al., 2003), and the adoption of Enterprise Resource Planning (ERP) systems (Law and Ngai, 2007). Third, an emerging body of research leverages unstructured data, particularly text, to quantify digital adoption. With the advancement of Natural Language Processing (NLP) and Large Language Models (LLM), economists can now analyze large-scale text data and extract valuable information (Cong et al., 2019). For instance, Chen and Srinivasan (2023) conduct a textual analysis on the annual reports of publicly listed firms in the U.S.. This approach involves creating a dictionary of digitization-related keywords and counting their occurrences within the reports. In our paper, we employ such textual analysis approach, specifically focusing on the context of firms listed on the Chinese stock market.

4.1.1 Baseline Measure of Digital Adoption

(i) Construction of A Dictionary of Digitization-Related Keywords

To begin with, we decide an initial set of keywords based on existing literature and political documents that are related to digital technology and digital economy. Particularly, we focus on the official document, Digital Economy and its Core Industries Statistical Classification (2021) released by the National Bureau of Statistics (NBS) of China[2]. In this classification system, the code 05 refers to industries with enhancement of efficiency through digitization. Under this category, for example, the first sub-category coded as 0501 specifically denotes *Smart Agriculture*. The note defines that Smart Agriculture involves the application of modern technology to the agricultural sector. Such technology includes remote sensing, geographic information systems (GIS), global positioning systems (GPS), Internet of Things (IoT), AI, big data, cloud computing, and unmanned aerial vehicles. These technologies are used to gather and analyze environmental data crucial for crop growth, such as soil properties, temperature, and humidity. The ultimate goal is to achieve accurate control and monitoring of crop cultivation and related activities. Such note on each digital-efficiency-enhancement-industry in this classification system specify the digital technologies that may be adopted in each industry. Therefore, this classification system can be served as a reference for the digital technology keywords selection. In addition, we gain insights from various types of official documents and plans related to digital transformation, as outlined by the government. For instance, the policy reports include annual reports presented by local governments outlining their work over the past year and to set forth plans and objectives for the coming year. One example is the Development Plan for Digital Transformation. It aims to guide the transition of traditional industries into digitally enabled industries. These documents collectively serve as a guidance for how to navigate digital transformation. Therefore, our digital keywords selection is developed based on those materials.

Subsequently, we employ NLP techniques to acquire a more comprehensive dictionary. Following

[2]See https://www.gov.cn/gongbao/content/2021/content_5625996.htm.

King et al. (2017), we expand the initial set of keywords into an extensive dictionary using the word embedding processes based on the well-established BERT model. By doing so, we can enrich our keywords dictionary, allowing for a contextually relevant and more comprehensive dictionary. The results based on such comprehensive dictionary serve as robustness checks. The two sets of digital-related keywords are presented in Table 1.

[Insert Table 1 here]

(ii) Keywords Occurrence and Refinement Algorithm

Using a Python-based web crawler, we systematically download annual reports from 2009 to 2022 from the Cninfo website [3], which is the information disclosure platform for listed firms designated by the China Securities Regulatory Commission. To ensure data integrity, only the latest or revised versions of these reports were retained for analysis. These reports are processed as textual inputs. Utilizing the jieba Python package, we perform text segmentation according to the two pre-defined dictionaries. Lastly, we aggregate the total occurrence counts and the scopes of keywords across all keywords by firm and year. In the main text, we focus on the total occurrence count, and the measure of digital scope is used for validity checks.

Admittedly, the raw count of keywords frequency in annual reports, while insightful, inherently contains many flaws. A common critique is its potential to capture "cheap talk" rather than substantive action. Firms, seeking to appeal to investor interests, might frequently discuss themes like digitization more as a gesture towards current trends than as a reflection of actual operational changes (Chen and Srinivasan, 2023). This phenomenon, where the discourse in annual reports may not necessarily translate into concrete practices, raises concerns about interpreting raw keyword counts as true evidence of digital adoption. This discrepancy between words and deeds, or cheap talk versus real walk, necessitates a more discerning algorithm for distinguishing.

To mitigate such limitation in the raw count measure, we propose an improved approach to capture

[3] See http://www.cninfo.com.cn/new/index.

substantive digital engagement. One possible improvement is to exclusively focus on the Management Discussion and Analysis (MD&A) section of the annual reports. Furthermore, after manually browsing six thousands of annual reports sentences containing the digital keywords, we identify and summarize some discernible patterns indicative of mere lip service as opposed to real digital transformation efforts. Based on this summary of patterns, we propose an algorithm of refinement for sentence-level examination to capture substantive digital engagement, as illustrated in Figure 1 and its note. This includes the presence of digital-related keywords coupled with additional conditions that signal tangible actions towards digitization. Such conditions include the mention of specific verbs that denote direct digital investment or deployment activities, financial terms that indicate quantifiable investment in digital technologies, and numerical references that suggest detailed accounting of digital initiatives. This methodology enables a more accurate assessment of a firm's digital transformation efforts, sifting out cheap talks about digitization, thereby enhancing the validity of the findings derived from keyword frequency analysis in the context of firm digital adoption.

[Insert Figure 1 here]

This detailed algorithm is as follows:

- Segment each annual report into individual sentences using delimiters '.', '?', '!' , and ';'.
- For each sentence, proceed as follows:
    - Initial Screening: If the sentence does not contain any digital-related keywords, discard it and terminate this iteration.
    - Further Condition: If the sentence has passed the initial screening (contains digital-related keywords), then assess it for additional criteria:
        - Action: Check if it includes words indicative of tangible actions or financial commitments about digital adoption.
        - Achievements: Determine whether it mentions accomplishments or outcomes related

to digital efforts.

- Numerical Evidence: Look for the presence of numbers, excluding the years from 2008 to 2023.

- Counting Keywords: If any of the conditions above (action, achievements, or numerical evidence) are met, sum the keywords count for this sentence. If none of these conditions are satisfied, conclude this iteration.

Last but not least, our final measure of digital transformation is a binary indicator for digital technology adoption based on the frequency measure of keyword occurrences. Specifically, if the firm-year level count is greater than zero, we define the adoption dummy as 1; otherwise, it is set to 0. Choosing between a binary adoption measure and a continuous frequency-based measure involves a trade-off between accuracy and informativeness. While frequency-based measures can offer more detailed information, they are prone to accuracy challenges, as noted by Cong et al. (2019). Since our primary objective is to classify firms into treatment and control groups based on adoption status, we opt for the binary measure, which is less subject to measurement error, following the approach of Acemoglu and Restrepo (2020) and Koch et al. (2021). Note that we conduct validity checks for the continuous intensity measure, as it contains more information in addition to the binary choice of adoption.

4.1.2 Validity Checks and Stylized Facts on Digital Adoption

We follow Sautner et al. (2023) to perform a battery of validity tests and show some stylized facts on firms' digital technology adoption.

(i) Effectiveness of the Refinement Algorithm[4]

As previously discussed, relying solely on the raw frequency count of keywords in annual reports may exaggerate the extent of digital adoption, attributable to the propensity of firms to overstate their digitization efforts. In response to this issue, we manually browse over 6,000 annual reports

[4]For brevity, we focus on the initial keywords dictionary as the baseline terms unless stated explicitly otherwise.

sentences containing the digital keywords, summarize the patterns of cheap talk and real walk, and propose a refinement algorithm, the procedure of which is detailed in the flowchart presented in Figure 1 and its notes. Here, we provide some excerpts illustrating the effectiveness of this algorithm in Table 2.

[Insert Table 2 here]

In Table 2, the first example from Ping An Bank's 2021 annual report highlights a substantial deployment of AI customer managers across over 900 scenarios, with a notable 149.3% year-over-year increase in customer service. The inclusion of these specific figures satisfies the "numeric evidence" criterion, reflecting a measurable commitment to digital initiatives. In contrast, the second excerpt from Ping An Bank's 2021 report is considered as a cheap talk, as it focuses on corporate vision and planning without offering concrete financial details or actions. Similarly, the message from Kweichow Moutai in 2022 is categorized as non-substantive. Although it discusses the acceleration of Smart Moutai as a digital strategy, these mentions remain within the realm of strategic intent, lacking evidence of tangible progress.

Figure 2 depicts the frequency counts of digital-related keywords in annual reports from 2009 to 2022. The blue bars represent the unrefined frequency of digital keywords in the entire annual report, showcasing the initial extent of self-reported digitization. The red bars apply the refinement algorithm to the full report, aiming to filter out instances of cheap talk where firms might have overstated their digital activities. And the yellow bars show the results of applying the same refinement algorithm, but only keep the MD&A sections of the annual reports, where the discourse is more focused on strategy and less on unsubstantiated mentions. There is a noticeable upward trend in counts of digital keywords from 2009 to 2022 across all measures. This suggests that firms have increasingly highlighted their digital initiatives in annual reports over the years. The raw count consistently surpasses the refined counts across all years by around two times. It underscores the discernment capability of the refinement algorithm, sifting out cheap talks to capture the substantive digital adoption. Such a refinement provides a more accurate assessment of a firm's

digital transformation.

[Insert Figure 2 here]

(ii) Face Validity of Digital Technology Terms

After getting some confidence in the refinement algorithm and the informativeness of the refined counts, we first study the patterns of specific digital technology terms. Figure 3 displays time-series of the refined counts of specific digital technology-related terms from 2009 to 2022. Each line tracks the average count of mentions of a particular digital term, averaged across all listed firms per year. These digital technologies show distinct trends. Cloud Computing has a robust upward trend, reflecting its growing importance in cloud strategies and operations. Artificial Intelligence presents a notable rise after 2015, indicating an increased emphasis on integrating intelligent systems into business processes. This surge aligns with the broader technological advancements and investments in AI research and applications. The initial trend of e-commerce shows an increase, reaching its peak around the year 2016. After this point, there is a decrease in the average number of counts over all listed firms. A possible reason for this decline probably because e-commerce has become a standard mode of operation for many businesses in China, the novelty may have worn off, leading to less discussion in annual reports. Internet of Things reflects a more recent but significant spike, which may be attributable to the advancement in sensors, connectivity, and data analytics technologies. Big Data experiences a considerable climb from 2013 to 2018, highlighting the critical role of data analytics in strategic decision-making.

[Insert Figure 3 here]

(iii) Time-Series Patterns with Different Keywords Dictionaries[5]

Firm's digitization level is measured at the report level, wherein each keyword individually contributes minimally to the aggregate digitization measures (Sautner et al., 2023). Nonetheless, in this

[5]We focus on the refined counts in this section.

part, we show more time-series patterns on the validity of the aggregate measures using different keywords dictionaries.

Figure 4 depicts the trends of the proportion of digital adopters from the year 2009 to 2022, using different dictionaries to identify digital-related content in annual reports. Digital adopter is defined as firms with digital count greater than zero. The blue line and red line are based on the occurrence of keywords in the initial digital technology dictionary. And the yellow line shows the proportion of digital adopters identified using the BERT-extended dictionary, which includes additional terms related to digital adoption to capture a broader range of digital-related content in the vector space. Moreover, the blue line is the occurrence searching through the full reports, while the other two lines are in MD&A sections. Figure 4 illustrates a clear upward trend in the proportion of digital adopters across all measures, especially a fast growth around the years 2012 to 2014 and 2017 afterwards. The extended dictionary captures a higher proportion of adopters than the initial dictionary, and we use this dictionary as an alternative robustness check for our treatment group construction and empirical results. The convergence of all measures in the more recent years might indicate a saturation point in digital adoption such as the e-commerce. Moreover, it is conceivable that this phenomenon arises from firms' initial emphasis on digital adoption activities in the adoption year, followed by a presumption of continued adoption and a decline in its explicit mention in subsequent years.

[Insert Figure 4 here]

(iv) Industry Variation

We provide industry-level evidence of digital adoption by calculating the mean and standard deviation of digital keywords count by the 4-digit WIND industry classification over the sample years 2009-2022. Table 3 indicates a heterogeneous distribution of digital adoption across industries. For instance, industries such as *Software Services*, *Telecommunication Services*, *Information Technology Hardware and Equipment*, and *Semiconductors and Semiconductor Equipment* exhibit notably high

mean values. This pattern aligns with the expectation that firms in these industries are inherently more engaged with digital technologies as they are digital technology producers. Moreover, financial sectors, such as *Insurance* and *Banking*, also exhibit high levels of digital adoption. In contrast, sectors like *Energy* and *Utilities* show lower mean counts, reflecting a slower rate of adoption and less intensive integration of digital technologies into their business modes.

The variation in standard deviation further underscores the diversity within industries concerning digital adoption. For example, industries such as *Insurance* and *Retailing* display high standard deviations, indicating that while some firms in those industries are deeply embedded with digital technologies, others may lag behind. Conversely, industries with lower standard deviations, such as *Pharmaceuticals, Biotechnology, and Life Sciences*, demonstrate a more uniform pattern of digital adoption.

[Insert Table 3 here]

Note that firms in *Software Services*, *Telecommunication Services*, *Information Technology Hardware and Equipment*, and *Semiconductors and Semiconductor Equipment* are digital technology producers, and they exhibit entirely different motivations and behaviors from our theoretical framework as discussed in Section 2.1. Therefore, these firms are excluded from the baseline empirical analysis in Section 5 but are included in the validity checks in Section 4.1.2.

To provide additional reference, we present digital reliance by sector in China as reported by Feng et al. (2025), calculated using Input-Output Tables from the NBS of China. Digital reliance is measured as the share of intermediate input sourced from the *Information Transmission, Software, and Information Technology Services* sector. Table 4 presents the average digital reliance of 42 sectors over the period 2009–2019. As expected, the *Information Transmission, Software, and Information Technology Services* sector itself exhibits the highest digital reliance, followed by sectors such as *Finance* and *Public Administration, Social Security, and Social Organizations*. This pattern aligns with our proxy for digital technology adoption, extracted from annual report text analysis.

Similarly, sectors like *Textiles*, *Food and Tobacco*, and *Petroleum, Coking Products, and Nuclear Fuel Processing Products* show the lowest digital reliance and are also positioned at the bottom in terms of digital adoption.

[Insert Table 4 here]

(v) Firm-Level Regression-Based Evidence

Our final validity tests examine the associations between our digital adoption intensity measures and proxies for digital technology engagement. These proxies are introduced in the context of convergent validity tests, where different methods of measuring the same underlying phenomenon are expected to show some degree of correlation (Sautner et al., 2023).

First, we use firm-level regressions to explore the association between several digital adoption intensity measures and the number of digital technology patents. Specifically, digital patents are identified by matching their International Patent Classification (IPC) with the Digital Economy and Core Industries Statistical Classification 2021. In Table 5, Panel A focuses on digital patent inventions, while Panel B considers all types [6] of digital patents.

The results in Table 5 indicate a consistent positive and significant relationship between digital adoption intensity measures and the digital patents across all columns. For example, the coefficient in Column (1) is 0.036, suggesting that a 1% increase in the raw count of digital adoption in full report is associated with a 3.6% increase in the number of digital patent inventions. Similar positive associations are observed for refined counts and digital scope in full report and MD&A sections, and by using the BERT-extended dictionary. These findings suggest that firms with higher digital adoption counts are more likely to be engaged in digital innovation activities, as reflected in their

[6] In China, patents are categorized into three types: invention, utility model, and design. Invention patents represent new technical solutions for a product, process, or improvement thereof. Utility model patents cover new technical solutions related to the shape, structure, or combination thereof of a product, suitable for practical use. Design patents focus on new designs that are aesthetic and applicable to industrial products. Invention patents are usually regarded as the most valuable type of patent. They involve the creation of new technical solutions or advancements in technology, often requiring a higher level of innovation and technical expertise compared to utility models and design patents. As a result, invention patents tend to have more rigorous examination processes and are often more highly valued in terms of intellectual property rights.

patent filings. These findings bolster the argument that our digital adoption intensity measures are valid indicators of a firm's engagement in digital technologies.

[Insert Table 5 here]

Second, the relationship between digital adoption measures and software assets (Panel A) and R&D expenditure over revenue (Panel B) is shown in Table 6. The results show a significant positive relationship association the digital adoption counts and the investment in software, although the sample size is smaller as not all firms report their software assets. Furthermore, as shown in Panel B of Table 6, there is a positive association between digital adoption and R&D expenditure ratio. This indicates that firms with higher digital adoption intensity measures are more likely to allocate resources towards research and development. Note that we do not distinguish the R&D expenditure in digital technology and other R&D fields. This could explain the less significant coefficient in front of the scope measure. Therefore, Table 6 further support the validity of our digital adoption measures, as they align well with the tangible investments firms make in digital technology.

[Insert Table 6 here]

Lastly, we investigate the link between digital adoption intensity measures and the amortization of intangible assets. Digital technology assets, such as software and data centers, typically have shorter amortization periods in China, compared with other types of intangible assets.[7] Therefore, a higher amortization rate can indicate a higher share on digital assets, given the same amount

[7]According to the Article 65 of the Implementation Regulations of the Corporate Income Tax Law, intangible assets are defined as non-monetary long-term assets held by enterprises for production, service provision, rental, or management purposes, including patents, trademarks, copyrights, land use rights, non-patented technology, and goodwill. Adoption of digital technologies such as cloud computing, data center renting, and AI service fall under these intangible assets and can be amortized accordingly. Furthermore, the Notice on Preferential Tax Policies for Corporate Income Tax (Cai Shui [2008] No. 1) stipulates that for enterprises purchasing software meeting the criteria for fixed or intangible assets, the amortization period can be shortened with approval from the tax authorities, with a minimum of 2 years. Additionally, the Notice on Further Encouraging the Development of the Software and Integrated Circuit Industries (Cai Shui [2012] No. 27) allows for a similar reduction in amortization periods for externally purchased software, which can be amortized over a minimum of 2 years without requiring tax department filing, provided that the enterprise maintains proper documentation.

of intangible assets. The results in Table 7 show a consistent positive relationship between our digital adoption measures and intangible asset amortization after controlling the intangible assets and other firm characteristics. This finding is enlightening, as it suggests that firms with higher digital adoption counts are indeed more involved in activities that lead to greater amortization of digital assets, providing further evidence of the validity of our intensity measures.

[Insert Table 7 here]

Overall, the regression-based results provide strong evidence that our digital adoption intensity measures are valid. The consistent correlations with diverse proxies of technological engagement suggest that our measures effectively capture firms' involvement in digital transformation.

4.2 Samples Construction

This section explains how we construct the treatment and control groups using our panel dataset. Based on the Statistical Classification of the Digital Economy and Its Core Industries (2021) by the NBS of China, we exclude firms operating in digital-related industries. These firms are producers of digital products and services and thus exhibit different motivations and behaviors compared with firms adopting digital technologies as discussed in 2.1. Specifically, firms classified under the 4-digit WIND industry classification of *Software Services*, *Telecommunication Services*, *Information Technology Hardware*, and *Semiconductors and Semiconductor Equipment* are excluded.

Let T_{it} be a dummy variable that equals 1 if the refined occurrence of digital keywords for firm i in year t is greater than zero, and 0 otherwise. It is an indicator of firm-year level digital adoption. We set a 5-year time horizon as a rolling window. The treatment group for batch t is defined as firms with no digital adoption in years $t-2$ and $t-1$, i.e., $T_{t-2} = T_{t-1} = 0$, and with digital adoption in years t, $t+1$, and $t+2$, i.e., $T_t = T_{t+1} = T_{t+2} = 1$. In contrast, the control group for batch t includes firms with no digital adoption in any of the years $t-2$, $t-1$, t, $t+1$, and $t+2$, i.e., $T_{t-2} = T_{t-1} = T_t = T_{t+1} = T_{t+2} = 0$.

Table 8 provides an illustration of the treatment group and control group definitions. Panel A presents the baseline setup using a 5-year horizon, while Panel B extends the horizon to 6 years for robustness checks. For each batch, a checkmark (✓) in year t indicates digital adoption ($T_{it} = 1$), and a cross (✗) indicates no adoption ($T_{it} = 0$). For example, in Panel A, the treatment group for batch 2011 includes all firms with no adoption in 2009 and 2010 ($T_{2009} = T_{2010} = 0$), and with digital adoption from 2011 to 2013 ($T_{2011} = T_{2012} = T_{2013} = 1$). Panel B extends the horizon to 6 years, including one more year of non-adoption before adoption.

[Insert Table 8 here]

For additional robustness checks, we extend the analysis to 7-year and 8-year horizons. In the 7-year horizon setup, we require three years of non-adoption before the treatment ($T_{t-3} = T_{t-2} = T_{t-1} = 0$) and four years of continuous adoption after ($T_t = T_{t+1} = T_{t+2} = T_{t+3} = 1$) for the treatment group. The corresponding control group consists of firms with no digital adoption throughout all seven years ($T_{t-3} = T_{t-2} = T_{t-1} = T_t = T_{t+1} = T_{t+2} = T_{t+3} = 0$). Similarly, the 8-year horizon extends the post-treatment period by one more year, requiring three years of non-adoption ($T_{t-3} = T_{t-2} = T_{t-1} = 0$) and five years of continuous adoption ($T_t = T_{t+1} = T_{t+2} = T_{t+3} = T_{t+4} = 1$) for the treatment group, with the control group showing no adoption across all eight years. These extended horizons help verify the persistence of digital adoption effects and ensure our findings are not sensitive to the time window specification.

In this manner, we establish a clear definition of the treatment and control groups. To verify accuracy, we conduct a manual random review of 1% of the annual reports, confirming that treatment firms adopted digital technology in the corresponding years while control firms did not.

To focus on firms with normal business operations, we restrict our sample to observations with positive earnings before interest, taxes, depreciation, and amortization (EBITDA) during the years t to $t+2$. In our baseline setup, which uses a 5-year horizon based on the refined adoption dummy, our final sample consists of 5,020 firm-batch observations from batch 2011 to batch 2020, with

the treatment group comprises 15.3% of the sample. Note that each observation contains a firm's information over the entire 5-year horizon rather than yearly observations.

4.3 Control Variables

In line with the existing literature, our control variables include firm age, size, cash ratio, leverage ratio, share of construction in progress, sales growth, sales growth volatility, proportion of high-skill workers, export dummy, and State-Owned Enterprise (SOE) dummy. We also consider industry characteristics such as the industry concentration, and the share of digitized firm within industry. Unless otherwise noted, data sources include the CEIC and WIND databases.

In detail, the definition of control variables are as follows. Firm age represents the number of years since the firm's establishment. Firm size is measured as the natural logarithm of the number of employees. Cash ratio is calculated as the ratio of cash and cash equivalents and financial assets held for trading to total assets. Leverage ratio is defined as the ratio of total liabilities to total assets. Sales growth is calculated as the annual growth rate of sales. We measure firm's risk by sales growth volatility (Griffin and Dugan, 2003). In particular, the sales volatility at year t is calculated as the standard deviation of sales growth in year t, $t-1$, and $t-2$. For the share of high-skilled workers, we process the CSMAR database, which provides information on the education information of firm employees. The share of high-skilled workers is then defined as the proportion of employees holding a diploma or higher. The SOE dummy equals to 1 if the firm is state-owned, and 0 otherwise. Similarly, the export dummy equals 1 if the firm has overseas revenues prior to its digital adoption; otherwise, it is 0. Note that this variable is set to 0 if the firm generates overseas revenues only after adopting digital technologies, as our primary interest lies in using this pre-condition variable to explain the digital adoption decision.

For the industry concentration and competition, we use the U.S. industry concentration because it not only reflects the competitive nature of the industry, but also is relatively exogenous in our

context. Data on this variable is sourced from the U.S. Census Bureau [8]. We collect the North American Industry Classification System (NAICS) 2-digit industry-level data on the shares of sales, value of shipments, or revenue of the top 50 largest firms in each industry for the years 2012 and 2017 and match these industries to the 4-digit WIND industry classification. Hence, the greater value is associated with higher concentration and less competition. Finally, the share of digitized firms within each 4-digit WIND industry classification is the mean of our binary measure of firm-level digital adoption at the corresponding industry-year level.

The definitions and summary statistics of key variables are reported in Table 9. We report data based on different timing, relative to the base year or batch t. For example, for each batch, we calculate the average values prior digital adoption, and label the subscription $-t$. For values in the base year t, we label the subscript t. In our main outcome models, we focus on the average values in years t to $t+2$, which are labeled the subscript $\bar{t}$. Finally, we aggregate all the batches and refer to the time relative to adoption only.

[Insert Table 9 here]

5 Empirical Results

In this section, we present the empirical estimations for the endogenous treatment effect model. It is important to note that while the results of binary choice model of digital technology adoption outlined in Equation (31) and the consequent firm outcome model in Equation (32) are introduced in separate sections, they are jointly estimated by MLE as simultaneous equations.

5.1 Probit Model: Key Determinants of Digital Adoption

Table 10 reports the results of the Probit model (31) over a 5-year time horizon, which spans 2 years before and 2 years after the year of digital adoption. In columns (1) and (2), both province-by-year

[8] See https://data.census.gov/table/ECNSIZE2017.EC1700SIZECONCEN?q=concentration&y=2017&n=N0600.00.

and industry-by-year FEs are included to account for regional and industrial time-varying factors. Columns (3) and (4) include only province-by-year FE [9], allowing us to examine the effects of industry-level concentration and the share of firms adopting digital technologies over time. Columns (1) and (3) report the estimation of coefficients in Probit model (31), while Columns (2) and (4) present the marginal effects derived from the Probit regressions.

[Insert Table 10 here]

The subscript t refers to the batch year, and $-t$ represents the average before digital adoption year t. Firm characteristics such as age, size, cash ratio, leverage, sales growth, the share of high-skilled workers, state ownership and export status are included in the form of their average values in year $t-2$ and $t-1$. This accounts for firm's decision-making process, which is typically influenced by information available 2-3 years before.

Table 10 highlights several key determinants of a firm's probability to adopt digital technologies. The findings indicate that, on average, larger firms with higher cash flow, lower leverage, fewer competing projects, greater sales growth with reduced volatility, and a higher proportion of skilled labour are more inclined toward digital adoption. Additionally, non-SOEs and exporting firms are more likely to embrace digital technologies. At the industry level, firms operating in sectors characterized by higher competition and a larger proportion of digital adopters show a greater propensity to adoption. Lastly, the instrumental variables are relevant as our expectation. Firms with longer MD&A sections, those situated in cities with higher broadband internet penetration, and those led by CEOs with IT-related backgrounds are significantly more likely to be classified as digital adopters.

We also present Probit model results under different time horizons in Table 11. Columns (1) and (2) use a 6-year horizon, including 3 years before, the adoption year, and 2 years after adoption.

[9] The numbers of observations in Columns (1) and (2) are smaller due to the Probit model's rich FE interactions. Specifically, certain FE groups exhibit no variation in the adoption dummy, leading to their exclusion and the omission of observations. For instance, in industry 5510 (Utilities) in the year 2014, there are 40 observations where the adoption dummy is consistently zero (i.e., no adopters). Consequently, this FE group is omitted from the Probit estimation, and the corresponding 40 observations are excluded.

Columns (3) and (4) extend the post-adoption period, covering 3 years before, the adoption year, and 3 years after adoption. Finally, Columns (5) and (6) focus on the 4-year period following adoption. The determinants of digital adoption observed in Table 11 align with those in our baseline Probit results in Table 10. This comparison suggests the robustness of our findings across different time horizons. While the sample sizes decrease as the time horizon extends due to the more stringent inclusion criteria, this does not affect our overall conclusions on the main determinants of digital adoption.

5.2 Outcome Model

5.2.1 Estimation Results

Table 12 presents the estimation results for the outcome model of profit after digital adoption. For each batch t, treatment is an adoption indicator that equals 1 if the firm adopts digital technology in years t, $t+1$, and $t+2$, and has not adopted in the preceding years $t-2$ and $t-1$, otherwise 0. The dependent variable is the three-year average of the natural logarithm of EBITDA post-adoption, i.e., the average value of the natural logarithm of EBITDA in years t, $t+1$, and $t+2$. Control variables are also calculated as three-year averages.

[Insert Table 12 here]

In Table 12, Columns (1) to (3) are estimated by MLE method, while Columns (4) to (6) are estimated by CF approach as described in Section 3.4. Both estimation approaches generate similar results, although MLE provides the most efficient estimation. At the bottom of Columns (1) to (3), ρ is the estimated correlation between the treatment-assignment error (v_i) and the outcome error (ε_i), and σ is the standard error of ε_i. In Columns (4) to (6), we report the estimation of β_2 in Equation (40), which shows the direction of selection bias.

Columns (1) and (4) of Table 12 present the results for the simple model of the outcome equation (37). The coefficients for the adoption dummy, γ, are significantly positive, indicating that

digital adoption is associated with improved EBITDA, even after accounting for observable firm characteristics and the endogenous nature of the adoption decision. As discussed in Section 3.3, in the simple model, $\gamma = ATE = ATET$. The estimates for the correlation between the adoption error and the outcome error from the MLE method, $\rho = \rho_0 = \rho_1$, and the self-selection control from the CF approach, $\beta_2 = \beta_{20} = \beta_{21}$, are all negative, suggesting that firms with unobservable characteristics that lead them to adopt digital technology tend to have lower profit, both in the adopting and non-adopting scenarios.

Columns (2) and (5) of Table 12 report the results for the outcome equation (33) of the control group, while Columns (3) and (6) present the results for the treatment group of the general model. At the bottom of Columns (1) to (3), the estimates for ρ_0, ρ_1, σ_0, and σ_1 are estimated, while in Columns (4) to (6), β_{20} and β_{21} are provided. The two different estimation approaches generate similar results. Moreover, the contribution of firm characteristics to profit are also similar between treatment and control groups, although some are insignificant in the treatment group due to less observations. The general model estimation provides us with the direction of self-selection: both ρ_0 and β_{20} are significantly negative, and both ρ_1 and β_{21} are insignificantly positive.

5.2.2 Decomposition: Treatment Effects and Selection Bias

In this section, we present the decomposition of treatment effect and selection bias of digital adoption on firm profit, based on the CF approach detailed in Section 3.4. We focus on discussing the results from the general model in detail, while the effects derived from the simple model are provided for comparison. As the simple model is a special case of the general model, differences in effect magnitudes are expected, but both models yield consistent empirical implications aligned with our theoretical framework.

We calculate the treatment effects and selection bias in Table 13, based on our CF estimation results in Columns (5) and (6) of Table 12. The observed averages for the control group and treatment group are $E[y_{i0}|A_i = 0] = 19.824$ and $E[y_{i1}|A_i = 1] = 20.124$, respectively. The counterfactuals

predicted by the general model show that if the control group had adopted digital technology, their logarithm of EBITDA would be $E[y_{i1}|A_i = 0] = 20.005$, and if the treatment group had not adopted, their log profit would be $E[y_{i0}|A_i = 1] = 19.454$. Additionally, using a 15.3% probability of digital adoption ($P = 0.153$), we estimate the potential outcomes for the entire population: $E[y_{i1}] = 20.023$ if all firms had adopted and $E[y_{i0}] = 19.768$ if none had adopted.

[Insert Table 13 here]

By definition, $ATE = E[y_{i1} - y_{i0}] = 0.255$, with standard deviation 0.436 obtained via bootstrapping. Such ATE equals the digital technology adoption cost of the marginal firm ($c_m = 2.55$) as discussed in Section 2.4. This seems to be a reasonable sum of $\delta = 10\%$, $f = 10\%$, and $s = 5.5\%$.

The observed $GMD = E[y_{i1}|A_i = 1] - E[y_{i0}|A_i = 0] = 0.300$, indicates that the mean profit of adopting firms is 30% higher than that of non-adopting firms. This GMD can be decomposed into two components: $ATET = E[y_{i1} - y_{i0}|A_i = 1] = 0.670$ and $SB = E[y_{i0}|A_i = 1] - E[y_{i0}|A_i = 0] = -0.370$. $ATET$ is of the key interest for firms that actually adopted digital technology. Our estimates suggest that had these firms not adopted, their logarithm of EBITDA would have been 19.454, compared to their actual logarithm of EBITDA of 20.124, yielding an ATET of 0.670, or a 3.33% profit gain relative to the counterfactual. In contrast, there are 84.7% of firms who actually did not adopt digital technology, and their $ATNT = E[y_{i1} - y_{i0}|A_i = 0] = 0.181$ but insignificantly different from zero.

We further provide a visual illustration based on the decomposition of general model results. Figure 5 plots the distributions of the non-adopting (upper panel) and adopting (lower panel) profits estimated by Columns (5) and (6) of Table 12. The black solid lines are the potential population means of the natural logarithm of EBITDA in two scenarios ($\mu_0 = E[y_{i0}] = 19.768$ and $\mu_1 = E[y_{i1}] = 20.023$), which suggest an improved potential mean profit with digital technology adoption, or $ATE = 0.255 > 0$.

[Insert Figure 5 here]

The red dashed lines indicate the treatment group means. In the panel of distribution of the non-adopting profit, the predicted potential mean of the *de facto* adopters ($E[y_{i0}|A_i = 1] = 19.454$) is less than average. However, in the panel of distribution of the adopting profit, the treatment group mean ($E[y_{i1}|A_i = 1] = 20.124$) exceeds the average. The difference between the two red dash lines is $ATET = 0.670$, which exceeds the observed GMD due to a negative self-selection.

Similarly, the green dashed lines are the control group means. The upper panel shows the control group mean $E[y_{i0}|A_i = 0] = 19.824$ is above average, while the lower panel of distribution of the adopting profit suggests the predicted potential mean of the *de facto* non-adopters ($E[y_{i1}|A_i = 0] = 20.005$) is slightly less than average. Therefore, $ATNT = 0.181$ is the difference between the two green dashed lines.

By definition (22) and (23), we compute $S_0 = E[y_{i0}|A_i = 1] - E[y_{i0}] = -0.314$ and $S_1 = E[y_{i1}|A_i = 1] - E[y_{i1}] = 0.101$. $S_0 < 0$ suggests that if all firms had no digital technology, firms that did adopt ($A_i = 1$) would, on average, lie to the left of the profit distribution, compared with the potentially non-adopting firms. Conversely, $S_1 > 0$ indicates that, with all access to digital technology, the adopters would outperformed the *de facto* non-adopters ($A_i = 0$). This is a pattern of "refugee sorting", as discussed in Section 2.3.

To understand the estimated patterns of $S_0 < 0$ and $S_1 > 0$, we trace back to Equations (24) and (25), which qualitatively reveal two key correlations: $\rho_{0c} > 0$ and $\rho_{1c} < 0$. Here, ρ_{0c} represents the correlation between unobservable factors affecting non-adopting profit (ε_{i0}) and adoption cost (ε_{ic}), while ρ_{1c} captures the correlation between factors influencing adopting profit (ε_{i1}) and adoption cost (ε_{ic}).

The observed "refugee sorting" pattern is driven by these correlations. Firms with higher non-adopting profits, which are intuitively the opportunity cost of digital adoption, tend to face higher adoption costs ($\rho_{0c} > 0$), while those with greater potential adopting profits typically encounter lower adoption barriers ($\rho_{1c} < 0$). SOEs exemplify this story: despite enjoying higher profits through policy support and market access advantages (Lin et al., 1998), they face substantial

digital transformation hurdles due to their complex organizational structures (Hu et al., 2024) and bureaucratic decision-making processes (Gu et al., 2020). These inherent characteristics, often difficult to quantify, contribute to SOEs' reluctance to embrace digital technologies, as evidenced by our Probit model estimates in Table 10.

In contrast, private enterprises operating in less privileged market positions tend to have greater adaptability. With lower opportunity costs and fewer institutional constraints, these firms often view digital transformation as both a transformative opportunity to outmaneuver competitors and a do-or-die effort. These enterprises, on average, succeed in leveraging digital technology to achieve significant profit gains.

Both $ATET$ and SB can be further decomposed into observable and unobservable components. The observable components are calculated as $(\boldsymbol{\beta}_1 - \boldsymbol{\beta}_0)' E[\boldsymbol{X}_i | A_i = 1]$ for $ATET$ and $\boldsymbol{\beta}_0' (E[\boldsymbol{X}_i | A_i = 1] - E[\boldsymbol{X}_i | A_i = 0])$ for SB. Consistent with existing literature, we find positive observables for $ATET$ and SB. The unobservable components, in contrast, depend on the estimates of β_{21} and β_{20}, as reported in Table 12. The unobservable factors in SB is significantly negative and dominant. To our knowledge, this paper is the first work to characterize the nature of self-selection in digital technology adoption, for both observable and unobservable factors. Our findings suggest that the significantly positive GMD of EBITDA is the result of a significantly positive and even higher $ATET$ and a negative unobserved SB.

5.2.3 Time-varying Effects of Digital Technology Adoption on Firm Profit

Table 14 presents the dynamic evolution of digital technology's effects on firm profits across different post-adoption periods. Panel A displays the results in the adoption year t, Panel B shows the effects one year after adoption, and Panel C reports the results two years post-adoption. For simplicity, we only report the estimates from the simple model using both MLE and CF approaches. Three-year average $E[y_{i0}]$ refers to the predicted three-year average value of the natural logarithm of non-adopting EBITDA.

[Insert Table 14 here]

The findings reveal several interesting patterns. First, the effect of digital adoption manifests immediately in the adoption year, with the effects ranging from 0.472 (MLE) to 0.612 (CF), suggesting that firms experience substantial profit gains shortly after adoption.

Second, the treatment effect strengthens in the first year post-adoption ($t+1$), reaching its peak with coefficients of 0.509 (MLE) and 0.671 (CF). This enhancement aligns with our "refugee sorting" narrative, where adopting firms - despite their initial disadvantages - successfully leverage digital technology to improve their profit position. The strengthening effect might reflect the completion of initial implementation phases and the realization of early efficiency gains.

Third, while remaining significantly positive, the effect moderates slightly in the second year post-adoption ($t+2$), with coefficients of 0.383 (MLE) and 0.500 (CF). This moderation should not be interpreted as diminishing returns but rather as a stabilization of benefits after the initial transformation phase. Compared to the predicted three-year average non-adoption profit ($E[y_{i0}]$) of approximately 19.8, these persistent positive effects represent substantial and sustained improvements in firm performance.

The temporal pattern of these effects offers important insights into the nature of digital transformation. Unlike traditional capital investments that might show delayed returns, digital adoption appears to generate immediate benefits that strengthen before stabilizing. This pattern suggests that firms in our "refugee sorting" framework not only overcome their initial disadvantages but also successfully embed digital capabilities into their operations, leading to sustained and persistent advantages.

6 Mechanisms

6.1 Demand Expansion and Cost Reduction

This section explores the two primary mechanisms through which digital adoption improves firm profits: demand expansion and cost reduction. We further analyze how these mechanisms operate differently across industry types, providing insights into the heterogeneous effects of digital transformation.

Digital technologies, particularly online marketing and digital advertising, can be powerful tools for expanding a firm's market reach. We examine the effects of digital adoption on four key variables: sales, market share, overseas revenue, and overseas revenue share. Sales is measured as the natural logarithm of operating revenue. Market share of firm i in industry j is defined as the proportion of firm i's operating revenue relative to the total operating revenue of all listed firms within the same industry j. Overseas revenue is calculated as the natural logarithm of revenue generated from overseas markets, with one added to account for zero values. Overseas revenue share is the proportion of total revenue derived from overseas markets.

Table 15 presents the estimation results for these demand-side outcomes. The results reveal substantial and increasing effects of digital adoption across all measures. Sales growth shows a persistent upward trend. This pattern suggests that firms become increasingly effective at leveraging digital technologies to generate revenue as they gain experience with these tools. Market share gains follow a similar trajectory, and these effects represent substantial relative gains, indicating that digital adopters effectively capture market share from non-adopting competitors. The effects on international market penetration, reflected by overseas revenue and its share, are also significantly positive. This is consistent with the existing findings in the literature that availability of digital technology enhances market access and in particular exporting, such as Hjort and Poulsen (2019). These results suggest that digital technologies are especially effective at helping firms expand market demand.

[Insert Table 15 here]

Digital technology adoption can also reduce costs. We examine three specific cost measures: costs of goods sold (COGS), labour costs, and management costs, all measured in terms of annual growth rates. Table 16 presents the estimation results for these cost-related outcomes. The temporal pattern of cost reduction effects differs notably from the demand expansion effects. COGS shows an immediate and substantial reduction in the adoption year, but this effect diminishes and becomes statistically insignificant in subsequent years. This pattern suggests that the introduction of digital technologies generate one-time efficiency gains in production processes.

Labour costs exhibit the largest immediate impact. Unlike COGS, the effect on labour costs persists, maintaining a 26.3% reduction two years post-adoption. This pattern is consistent with successful automation and digital transformation of labour-intensive processes, suggesting a structural change in firms' labour requirements. Management costs show a more modest pattern, with an 18.5% reduction in the adoption year that becomes statistically insignificant in subsequent years. This temporal pattern may reflect initial efficiency gains from digital systems offset by the ongoing costs of maintaining digital infrastructure.

[Insert Table 16 here]

To better understand how these mechanisms operate across different contexts, we conduct separate exercises for demand-driven and production-driven industries. Based on WIND 4-digit industry classification codes, we identify production-driven industries as those in Energy (1010), Materials (1510), Capital Goods (2010), Automobiles and Components (2510), Health Care Equipment and Services (3510), and Pharmaceuticals, Biotechnology, and Life Sciences (3520). All remaining industries are classified as demand-driven.

Table 17 presents these heterogeneous effects. The results reveal substantial differences in how digital technologies affect different types of industries. Column (2) suggests that demand-driven

industries show larger sales increases. In contrast, production-driven industries in Column (3) show more modest sales growth. [10] This gap suggests that digital technologies are particularly effective for customer-facing operations and market expansion in demand-driven industries.

The cost reduction patterns also differ markedly across industry types. Production-driven industries show more consistent cost reductions in Column (6), with a significant -30.2% effect in the adoption year that maintains at -16.4% by year $t+2$. In contrast, Column (5) indicates that demand-driven industries show insignificant cost reduction.

These findings reveal important nuances in how digital adoption creates value. The benefits appear to operate through two distinct pathways: revenue enhancement through market expansion and customer reach, and operational efficiency through cost reduction. The relative importance of these pathways varies by industry type and time horizon. Revenue benefits tend to be progressive and increasing, while cost benefits are often immediate but may partially diminish over time.

[Insert Table 17 here]

6.2 Productivity Gain and Capital Labour Substitution

We next examine how digital adoption affects firm productivity and factor intensity. These real-side effects are particularly important for understanding the economic mechanisms through which digital technologies transform production processes.

We calculate productivity by two key measures. First, we compute Total Factor Productivity (TFP) using three widely-recognized approaches: the Olley-Pakes (OP) approach (Olley and Pakes, 1992), the Levinsohn-Petrin (LP) method (Levinsohn and Petrin, 2003), and the generalized method of moments (GMM) estimation following Blundell and Bond (2000). These methods address various

[10] The sum of observations in Columns (2) and (3) does not equal that of Column (1) because, in the Probit model with rich fixed-effects (FE) interactions, some FE groups exhibit no variation in the adoption dummy. For instance, in province 62 in 2012, all four observations had a zero adoption dummy, resulting in the exclusion of this FE group and the omission of these observations. As a result, the more granular the subgroup, the more observations are omitted, leading to a smaller total sample size.

empirical challenges in productivity estimation, including simultaneity bias and selection issues. Therefore, we report all of them as robustness checks. Second, we calculate labour productivity as the natural logarithm of operating revenue per employee.

Table 18 presents the estimation results for productivity enhancement. The effects are both economically and statistically significant across all measures. The OP TFP estimate shows an immediate increase in the adoption year with a rising pattern by year $t+2$. The LP and GMM estimates reveal similar patterns but with larger magnitudes. The consistency across different estimation approaches of TFP strengthens the robustness of our findings. Labour productivity exhibits comparable improvements. The progressive nature of these productivity gains suggests that firms undergo a learning process in effectively integrating digital technologies into their operations. The persistence and growth of these effects indicate that productivity enhancement represents a structural improvement rather than a temporary boost.

[Insert Table 18 here]

Digital technology adoption may fundamentally alter firm's optimal input combination, particularly through capital-labour substitution. We examine this argument by analyzing the impacts of digital adoption on intangible capital, amortization, and employment growth. In literature, evidence from the U.S. market indicates that automation technologies often replace certain labour functions, resulting in shifts in labour demand (Acemoglu et al., 2022; Acemoglu and Restrepo, 2020; Bessen, 2018). Consistent with this literature, Table 19 presents similar results.

[Insert Table 19 here]

The evidence strongly supports substantial capital deepening following digital adoption. In Column (1), intangible capital shows an immediate increase in the adoption year, peaking in year $t+1$ before slightly moderating in year $t+2$. As discussed in Section 4.1.2, Column (2) uses the natural logarithm of amortization as an alternative proxy. The increase in amortization further confirms the

substantial scale of digital capital investment. Concurrent with this capital deepening, we observe significant labour substitution. In Column (3), Employment growth decreases in the adoption year, with the effect moderating to -13.6% in year $t+1$ and becoming statistically insignificant by year $t+2$.

The joint patterns in capital investment and employment suggest that digital adoption induces a significant shift toward capital-intensive production patterns. The front-loaded nature of both effects—strongest in the adoption year and moderating thereafter—indicates that this transformation represents a one-time structural change in factor combination.

These findings have important implications for understanding the aggregate effects of digital transformation. The combination of productivity enhancement and capital-labour substitution suggests that digital technologies may contribute to both economic growth and labour market polarization. The productivity gains indicate potential aggregate efficiency improvements, while the capital-labour substitution patterns suggest possible distributional consequences across different types of labour.

6.3 Valuation and Profitability

The final dimension of our analysis examines how digital adoption affects market valuation and firm profitability. We first examine market valuation effects through Tobin's Q, which captures the market's assessment of firm value relative to replacement cost. In Column (1) of Table 20, we observe a significant positive and immediate market response to digital adoption announcements, with Tobin's Q is on average 0.690 units higher than non-adopters in the adoption year. However, this market response appears to be transient. The effect diminishes to 0.476 in year $t+1$ and 0.380 in year $t+2$, with both estimates becoming statistically insignificant. The temporal pattern is consistent with rational expectations, where markets efficiently incorporate new information about the challenges and realized benefits of digital transformation.

Our theoretical framework assumes that individual firm adoption decisions do not significantly affect

industry equilibrium, implying a constant conduct parameter θ, and therefore we do not have an a prior hypothesis regarding the effects of digital adoption on industry competition as captured by the conduct parameter θ. This assumption warrants empirical validation, as changes in competitive intensity could affect firm markup and profitability.

Therefore, we examine profitability as an inverse proxy for markup to assess the competitive equilibrium in industries after the adoption of digital technologies. This analysis helps infer how digitization may alter market power and competitive dynamics. In Column (2) of Table 20, we find no significant effect on profitability in the adoption year or one year post-adoption, with point estimates of -0.051 and -0.047, respectively, although by year $t+2$, we observe a marginally significant negative effect of -0.084. These results have important implications for our theoretical framework.

[Insert Table 20 here]

First, the lack of significant profitability effects in the initial periods supports our assumption that individual adoption decision does not substantially alter industry competitiveness in the short run. This validates our theoretical approach of treating the conduct parameter θ as constant immediately after adoption.

Second, the emergence of a negative profitability effect by year $t+2$ suggests that competitive forces may eventually erode and dilute the benefits of digital adoption, followed by more firms entering the digital market. This pattern is consistent with models of technology diffusion where early adopters gain temporary advantages that dissipate as competitors catch up. The negative point estimate could reflect both increased competitive pressure and the costs of maintaining digital infrastructure.

These findings help complete our understanding of the digital adoption mechanism. While we document substantial improvements in productivity, market reach, and operational efficiency, the lack of sustained profitability gains suggests that these benefits are ultimately competed away.

The results also highlight an important distinction between private and social returns to digital

adoption. While individual firms may not capture sustained profitability gains, the productivity and efficiency improvements we document suggest substantial social benefits from digital transformation, suggesting an improved consumer surplus since the benefits appear to flow primarily to consumers. This divergence between private and social returns may have important implications for optimal technology adoption policies under positive externality.

7 Conclusion

This paper investigates the endogenous selection and causal effects of digital technology adoption using a theoretical framework and its empirical specification. We develop a general theoretical framework that describes firm's adoption decision based on potential benefits, costs, and unobservable factors. It allows for correlation between non-adopting profit, adopting profit, and adoption cost. Our empirical strategy employs an endogenous treatment effect model that accounts for self-selection of digital adoption. Within this endogenous treatment effect model, we identify causal effects with three instruments: annual report length, city-level broadband penetration, and CEO IT background. This framework also allows us to disentangle selection bias from causal effect while accounting for both observable and unobservable factors influencing adoption decision. To empirically test these relationships, we obtain a rigorous indicator of digital adoption through textual analysis based on annual reports of Chinese public firms from 2009 to 2022.

Our analysis yields several important findings that contribute to understanding both the determinants and consequences of digital transformation in the modern economy.

First, we uncover a novel "refugee sorting" pattern in digital adoption. While observable firm characteristics suggest a positive selection consistent with existing literature, unobservable factors reveal a significant and dominant negative selection. Firms that adopt digital technologies would have underperformed in the non-adopting scenario but achieve superior performance post-adoption.

Second, we identify the key factor driving this pattern: a positive correlation between non-adopting

profit potential and adoption cost. This correlation creates a barrier that particularly affects firms with strong existing capabilities, as they face higher opportunity costs of disrupting their current operations. This insight helps explain the digital transformation hesitancy observed among some SOEs.

Third, our analysis reveals the various benefits of digital adoption serving as channels of profit gain. Adopting firms experience substantial demand expansion, evidenced by increased sales, market share, and growth in overseas revenue and its share. They also achieve significant cost reductions across multiple dimensions, with immediate effects on COGS, labour costs, and management costs. Furthermore, productivity improvements are substantial and persistent. We also document significant capital-labour substitution effects. This finding has important implications for understanding how digital technologies reshape firm's production patterns and employment markets.

Finally, our analysis of market valuation and profitability reveals an interesting temporal pattern. While markets initially respond positively to digital adoption announcements, these valuation gains prove to be transient. Moreover, the absence of sustained profitability improvements, despite clear operational benefits, suggests that competitive forces eventually dissipate firm-specific advantages from digital adoption and that the benefits, while substantial, may be primarily captured by consumers rather than firms in the long run.

These findings have important implications for public policy. Our paper implies the distinction between private and social returns to digital adoption. While individual firms may not capture sustained profitability gains in the long run, the substantial improvements in productivity, market access, and operational efficiency suggest significant social benefits from digital transformation. This divergence between private and social returns might justify policy interventions to encourage digital adoption, particularly among firms facing high adoption costs despite potential benefits.

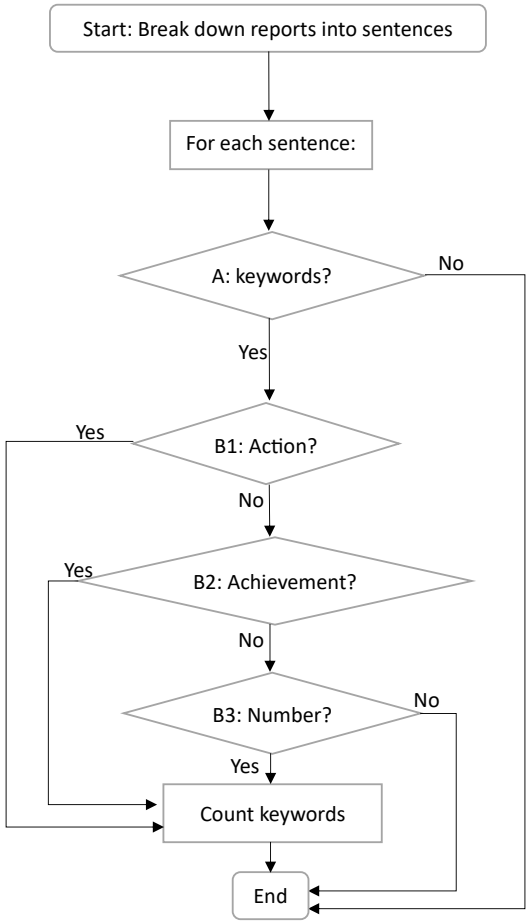


Figure 1: Refinement Algorithm of Digital Technology-Related Keywords Count

Note: The flowchart describes the refinement algorithm proposed to examine sentences in annual reports for evidence of substantive digital adoption by firms. A sentence is retained for keywords counting if it meets a specific set of criteria: (A) It contains any digital-related keywords, and concurrently, (B) it also includes references that signify actual financial commitment or tangible actions towards digitization. These references are categorized into three subsets: (B1) the presence of verbs such as *expenditure*, *cost*, *investment*, *purchase*, *adoption*, *procurement*, *research and development*, *equipment*, *order*, *preorder*, *prepayment*, *establishment*, indicating direct actions or financial allocations towards digital initiatives; (B2) the appearance of financial units or indicators like *yuan*, *thousand*, *million*, *billion*, *awards*, *%*, signifying quantifiable achievements, accomplishments, or outcomes related to digital efforts; and (B3) the mention of any numbers excluding the years 2008-2023 to avoid confusion with year references.

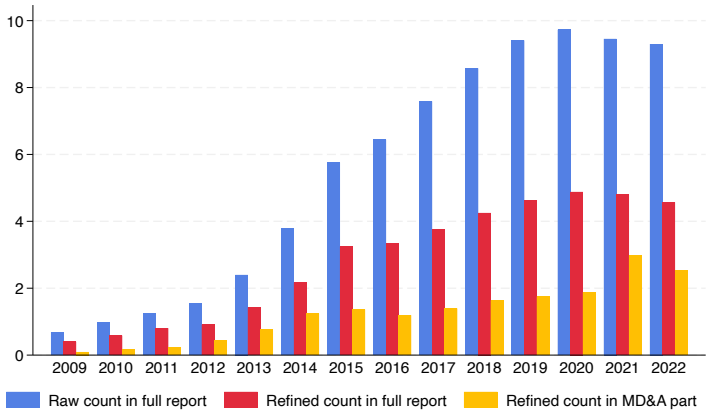


Figure 2: Trends in Digital Technology-Related Keywords Counts

Note: This figure plots the trends of frequency counts of digital-related keywords in annual reports from 2009 to 2022. The blue bars represent the unrefined frequency counts of digital keywords in the entire annual report. The red bars apply the refinement algorithm to the full report, aiming to filter out instances of cheap talk where firms might have overstated their digital activities. And the yellow bars show the results of applying the same refinement algorithm, but only keep the MD&A sections of the annual reports, where the discourse is more focused on strategy and less on unsubstantiated mentions.

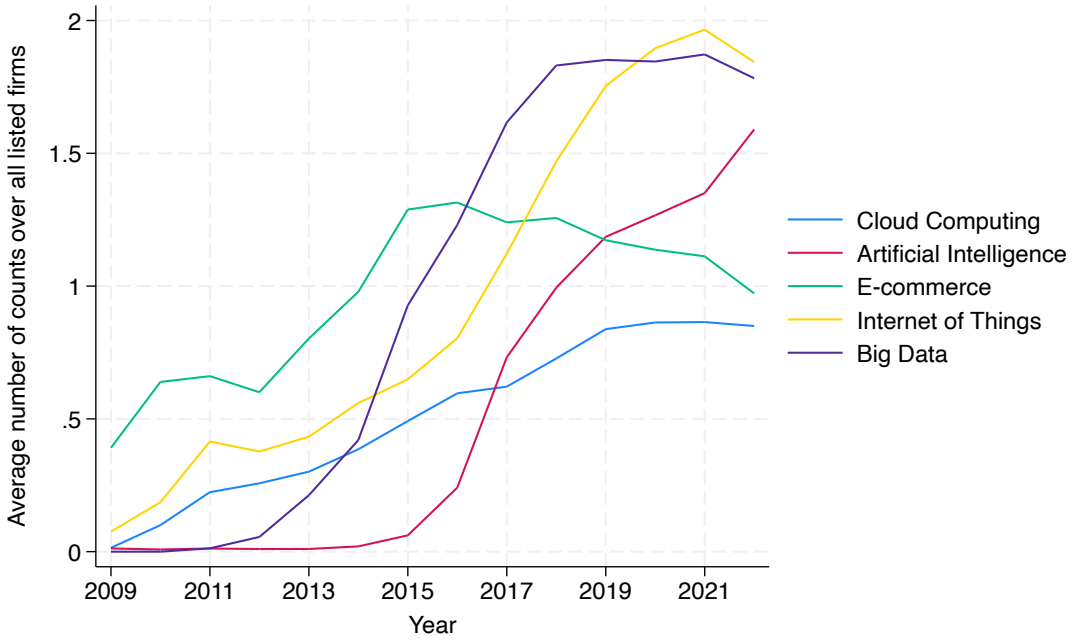


Figure 3: Trends in Digital Technology Adoption

Note: This figure shows different trends of the average refined occurrences of digital technology terms from 2009 to 2020. Each line tracks the evolution of a term's occurrence in annual reports, averaged across all firms per year. The terms are Cloud Computing, Artificial Intelligence, E-commerce, Internet of Things, and Big Data.

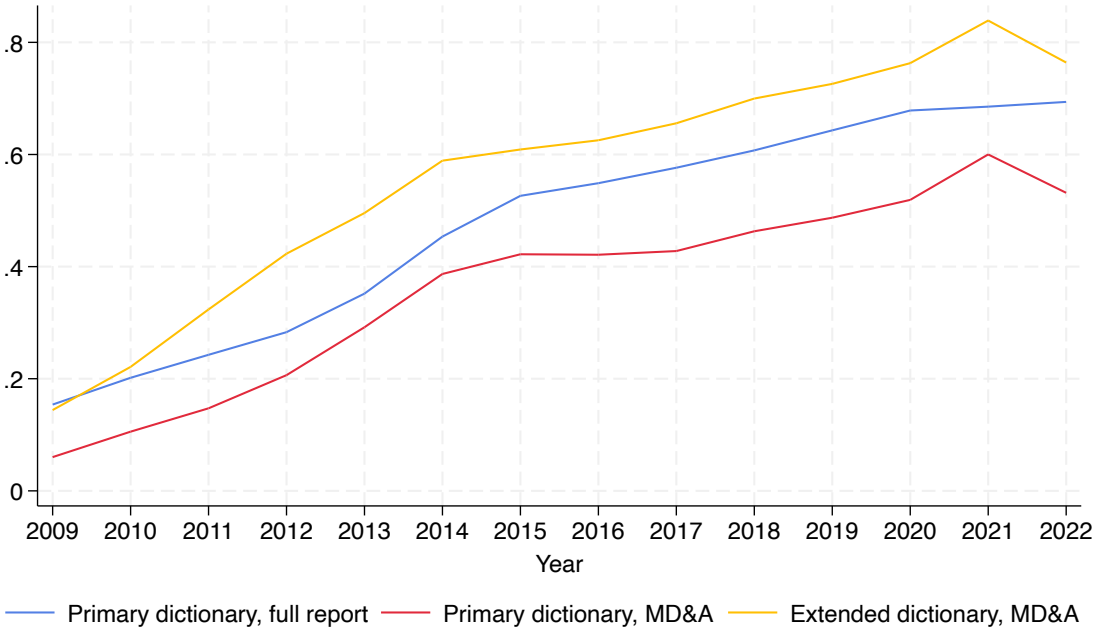


Figure 4: Trends in the Proportion of Digital Adopters

Note: This figure shows the trends of the proportion of digital adopters from 2009 to 2022, using different dictionaries to identify digital-related contents in annual reports. Digital adopter is defined as firms with digital count greater than zero. The blue line and red line are based on the occurrence of keywords in the initial digital technology dictionary. And the yellow line shows the proportion of digital adopters identified using a BERT-extended dictionary, which includes additional terms related to digital adoption to capture a broader range of digital-related content in the vector space. Moreover, the blue line is the occurrence searching through the full reports, while the other two lines are in MD&A sections.

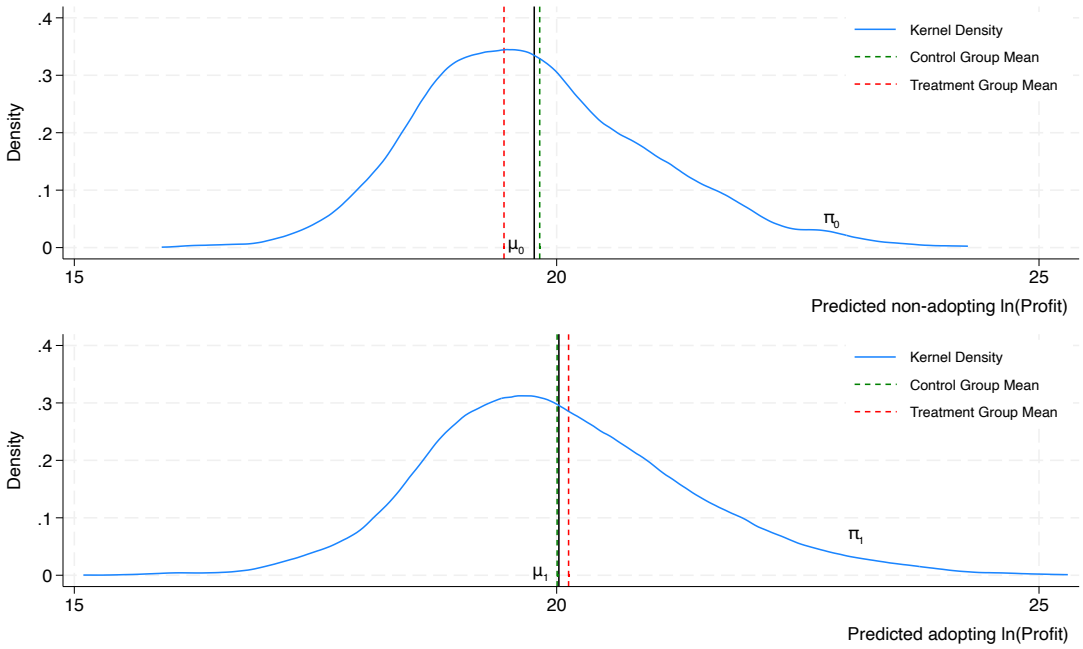


Figure 5: Distribution of Non-Adopting and Adopting Profits

Note: This figure plots the distributions of the natural logarithm of non-adopting EBITDA (upper panel) and the natural logarithm of digital-adopting EBITDA (lower panel) as estimated by the CF approach from the general model. The black solid lines represent the potential population means of the natural logarithm of EBITDA. The red dashed lines indicate the treatment group means, while the green dashed lines are the control group means.

Table 1: Dictionaries of Digitization-Related Keywords

Initial keywords dictionary (78 terms)	Artificial Intelligence, Augmented Reality, Autonomous Driving, B2B, B2C, Billion-Level Concurrency, Big Data, Blockchain, Brain-Inspired Computing, Business Intelligence, C2B, C2C, Cloud Computing, Cognitive Computing, Converged Architecture, Credit Reporting, Cyber-Physical Systems, Data Mining, Data Visualization, Deep Learning, Differential Privacy Technology, Digital Currency, Digital Finance, Digital Marketing, E-Commerce, Exabyte-Level Storage, Facial Recognition, Fintech, Geographic Information Systems (GIS), Global Positioning Systems (GPS), Graph Computing, Green Computing, Heterogeneous Data, Identity Verification, Image Understanding, In-Memory Computing, Industrial Internet, Intelligent Data Analysis, Intelligent Robots, Internet Finance, Internet Healthcare, Internet of Things, Investment Decision Support Systems, Machine Learning, Mixed Reality, Mobile Connectivity, Mobile Internet, Mobile Payments, Multi-Party Secure Computing, Natural Language Processing, Network Connectivity, NFC Payments, O2O, Open Banking, Quantitative Finance, Remote Sensing, Semantic Search, Smart Agriculture, Smart Customer Service, Smart Energy, Smart Environmental Protection, Smart Financial Contracts, Smart Grid, Smart Healthcare, Smart Home, Smart Investment Advisory, Smart Marketing, Smart Tourism and Culture, Smart Transportation, Smart Wearables, Stream Computing, Text Mining, Third-Party Payments, Unmanned Aerial Vehicles, Unmanned Retail, Virtual Reality, Voice Recognition
BERT-extended dictionary (303 terms)	5G, 5G Technology, AI, AI-Human Interaction, Algorithms, Algorithm Models, Applications, Artificial Intelligence, Association Analysis, Auto-Capture, Automation, Augmented Reality, Autonomous Driving, B2B, B2C, Behavior Analysis, Big Data, Big Data Center, Big Data Mining, Big Data Models, Big Data Platform, Big Data R&D, Big Data Risk Control, Big Data Services, Big Data Technology, Billion-Level Concurrency, Biometric Technology, Blockchain, Brain-Inspired Computing, Business Intelligence, C2B, C2C, Central Processing Unit, ChatGPT, Client Software, Cloud, Cloud Applications, Cloud Architecture, Cloud Adoption, Cloud Computing, Cloud Computing Platform, Cloud-Based, Cloud-Network Integration, Cloud Payments, Cloud Platform, Cloud Service Platform, Cloud Service Platforms, Cloud Services, Cloud Systems, Cognitive Computing, Consensus Mechanisms, Converged Architecture, Credit Reporting, Customer Profiling, Cyber-Physical Systems, Cybersecurity, Data Analyst, Data Center, Data Center Clusters, Data-Driven, Data Elements, Data Encryption, Data Governance, Data Integration, Data Labels, Data Management, Data Mining, Data Mining Algorithms, Data Models, Data Platform, Data Technology, Data Visualization, Data Warehouse, Deep Learning, DevOps, Differential Privacy Technology, Digital Banks, Digital Currency, Digital Economy, Digital Ecosystem, Digital Factory, Digital Finance, Digital Infrastructure, Digital Marketing, Digital Office, Digital Operations, Digital Payments, Digital Projects, Digital Supervision, Digital Technology, Digital Transformation, Digital Twins, Digitalization, Digitalized Business, Distributed Architecture, Distributed Computing, Distributed Databases, Distributed Ledger, Distributed Storage, E-Commerce, E-Signatures, E-Social Security Cards, Electronic Banking, Electronic Channels, Electronic Finance, Electronic Payments, Encrypted Storage, Exabyte-Level Storage, Facial Payment, Facial Recognition, Financial Cloud, Financial Digitalization, Fintech, Geographic Information Systems (GIS), Global Positioning Systems (GPS), Graph Computing, Gray-Scale Deployment, Green Computing, Heterogeneous Data, High Bandwidth, Human-Machine Collaboration, ICT Infrastructure, Identity Authentication, Identity Recognition, Identity Verification, Image Recognition, Image Understanding, Immersive Technology, In-Memory Computing, Inclusive Finance, Industrial Internet, Information Infrastructure, Information Security, Information Technology, Innovation Labs, Inspection Robots, Integrated Circuit Cards, Integrated Circuit Chips, Intelligence, Intelligent Advisory, Intelligent Analytics, Intelligent Auditing, Intelligent Customer Service, Intelligent Data Analysis, Intelligent Economy, Intelligent Investment Advisory, Intelligent Recommendations, Intelligent Recognition, Intelligent Robots, Intelligent Search, Intelligent Sensing, Intelligent Systems, Intelligent Technology, Internet, Internet Finance, Internet Healthcare, Internet of Things, Investment Decision Support Systems, IT Infrastructure, Knowledge Graph, Large-Scale Data Processing, Lifecycle Management, Machine Learning, Metaverse, Mixed Reality, Mobile Applications, Mobile Banking, Mobile Connectivity, Mobile E-Commerce, Mobile Finance, Mobile Internet, Mobile IoT, Mobile Payments, Mobile Terminals, Model Computing, Model Development, Multi-Party Secure Computing, Multimedia Data, Narrowband IoT, Natural Language Processing, Near-Field Communication, Network Banking, Network Communication, Network Finance, Network Identity Authentication, Networked Connectivity, Neural Networks, NFC Payments, NFTs, NLP, O2O, Online, Online Banking, Online Finance, Online Payments, Online Services, Online Transactions, Open Banking, Open Platform, OpenAI, POS, Private Cloud, Quantum, Quantum Communication, Quantum Technology, QR Code Devices, Quantitative Finance, Real-Time Monitoring, Remote Sensing, RFID, Robots, Satellite, Satellite Remote Sensing, Scenario-Based Marketing, Secure Multi-Party Computing, Semantic Search, Smart Agriculture, Smart Algorithms, Smart Banking, Smart Branches, Smart Customer Service, Smart Energy, Smart Environmental Protection, Smart Finance, Smart Grid, Smart Healthcare, Smart Home, Smart Marketing, Smart Risk Control, Smart Tourism, Smart Transportation, Smart Wearables, Software, Software Development Centers, Software Development Kits, Software-Defined Networks, Stream Computing, Telecom Infrastructure, Text Mining, Third-Party Payments, Trusted Computing, Unmanned Aerial Vehicles, Unmanned Retail, Virtual Reality, Virtualization, Visual Technology, Voice Recognition, Voiceprint Recognition

Table 2: Examples of Distinguishing Substantive Digital Adoption in Annual Reports

Firm	Industry	Year	Excerpt containing digital keyword	Remark
Ping An Bank	Banking	2021	By the end of 2021, AI customer managers had been deployed in over 900 scenarios, with the average monthly service to customers increasing by 149.3% year-over-year.	The algorithm identifies the real adoption of AI due to the numbers, satisfying criteria B2 and B3.
Ping An Bank	Banking	2021	We will enhance the application of leading technologies, deepen the innovative application of technologies such as IoT, blockchain, AI, and big data in the financial sector, and promote the digital and intelligent upgrade of business management.	This excerpt is a cheap talk as it does not fulfill any of the three criteria (B1 to B3), instead discussing corporate vision and planning without tangible commitments.
GREE Electric Appliances	Manufacture	2022	The company, through long-term big data analysis of the operating characteristics of domestic multi-connected machines and combined with research on ..., and the compressor's APF efficiency value has increased by 8%, its low-load operation has improved by an additional 20%, and the speed of cooling and heating upon startup has doubled.	Real adoption of big data is recognized by the refinement algorithm due to the presence of numbers 8% and 20%, meeting criteria B2.
Kweichow Moutai	Liquor	2022	The third strategy is to accelerate the construction of Smart Moutai, vigorously implement digital empowerment, comprehensively advance the development of three major sectors: smart industrial chains, smart parks, and the industrial internet. This initiative aims to promote the establishment of an industrial internet platform, further nurture and develop new drivers of growth, and propel digital transformation to new heights.	This excerpt is labeled as a cheap talk as it does not fulfill any of the three criteria, instead discussing corporate vision and planning without tangible commitments.
GREE Electric Appliances	Manufacture	2014	The Gree mobile phone is set to be launched soon. As the carrier responsible for linking and controlling smart home appliance products, Gree aims to be at the forefront of the smart home entrance battle. The ultimate goal is to achieve remote and visualized control. In the future, it also has the potential to enter the field of big data services.	The statement is a cheap talk since it lacks any of the criteria required and only talks about future plan without tangible commitments or numerical indicators.

Note: This table presents some excerpts containing digital-related keywords. The criteria B1, B2, and B3 are illustrated in the notes of Figure 1.

Table 3: Summary Statistics of Industry-level Digital Keywords Counts

Code	Industry Name	Observations	Mean	Std. Dev.
4510	**Software Services**	**2407**	**33.25**	**38.80**
5010	**Telecommunication Services**	**45**	**20.40**	**18.97**
4030	Insurance	69	14.51	29.45
4520	**Information Technology Hardware**	**3417**	**13.88**	**22.87**
4010	Banks	322	12.87	13.85
3010	Food and Staples Retailing	213	9.94	10.48
2540	Media	898	9.57	13.08
4530	**Semiconductors and Semiconductor Equipment**	**904**	**7.97**	**17.39**
2550	Retailing	698	7.41	28.33
4020	Diversified Financials	681	6.13	10.24
2520	Consumer Durables and Apparel	2000	5.28	8.91
3510	Health Care Equipment and Services	836	5.12	9.04
2010	Capital Goods	7955	4.26	10.86
2030	Transportation	1209	3.99	14.64
2020	Commercial and Professional Services	1040	3.77	8.97
2530	Consumer Services	510	3.57	7.05
3030	Household and Personal Products	127	3.09	4.93
6010	Real Estate	1300	2.73	12.04
2510	Automobiles and Components	1710	2.32	8.35
3020	Food, Beverage and Tobacco	1843	2.21	4.41
3520	Pharmaceuticals, Biotechnology and Life Sciences	2728	1.17	2.88
1510	Materials	6380	1.14	3.24
1010	Energy	839	0.93	3.05
5510	Utilities	1343	0.88	2.17

Note: This table shows the mean and standard deviation of digital keywords count by the 4-digit WIND industry sector in 2009-2022. The industries in bold (Software and Services, Telecommunication Services, Information Technology Hardware and Equipment, and Semiconductors and Semiconductor Equipment) are considered as digital technology producers. Firms in such industries are excluded from the baseline empirical analysis in Section 5 but are included in the validity checks in Section 4.1.2.

Table 4: Digital Reliance across Sectors in China

Code	Sector Name	Digital Reliance
29	Transportation, warehousing, and postal services	17.37%
32	Finance	9.11%
42	Public administration, social security, and social organizations	7.62%
28	Wholesale and retail	4.85%
39	Education	4.32%
41	Culture, sports, and entertainment	3.59%
30	Accommodation and dining	2.62%
26	Water production and supply	2.55%
25	Gas production and supply	2.19%
40	Health and social work	2.06%
37	Water conservancy, environment, and public facility management	2.00%
36	Comprehensive technical services	1.97%
33	Real estate	1.93%
27	Construction	1.86%
5	Non-metallic minerals and other mining products	1.66%
35	Research and experimental development	1.35%
34	Rental and business services	1.33%
19	Electrical machinery and equipment	1.16%
38	Resident services, repair, and other services	1.06%
15	Metal products	0.97%
31	Information transmission, software, and information technology services	0.89%
18	Transportation equipment	0.78%
2	Coal mining and selection products	0.77%
20	Communication equipment, computers, and other electronic devices	0.77%
21	Instruments and meters	0.76%
16	General equipment	0.76%
23	Metal products, machinery, and equipment repair services	0.73%
3	Petroleum and natural gas extraction products	0.73%
8	Textile apparel, footwear, headgear, leather, down, and their products	0.70%
24	Electricity and heat production and supply	0.65%
9	Wood processing products and furniture	0.58%
13	Non-metallic mineral products	0.55%
14	Metal smelting and rolling products	0.54%
4	Metal ore mining and selection products	0.53%
17	Specialized equipment	0.53%
12	Chemical products	0.48%
1	Agricultural, forestry, animal husbandry, and fishery products and services	0.46%
22	Other manufacturing products and waste materials	0.41%
10	Paper, printing, and educational and sports goods	0.39%
7	Textiles	0.33%
6	Food and tobacco	0.31%
11	Petroleum, coking products, and nuclear fuel processing products	0.25%

Note: This table shows the average digital reliance of 42 sectors based on Input-Output Tables from the NBS of China over the period 2009–2019. Digital reliance is measured as the share of intermediate input sourced from the *Information Transmission, Software, and Information Technology Services* sector (Feng et al., 2025).

Table 5: Validity Checks: Digital Technology-Related Patents

	(1) ln(raw count)	(2) ln(refined count)	(3) ln(digital scope)	(4) ln(raw count in MD&A)	(5) ln(refined count in MD&A)	(6) ln(scope in MD&A)	(7) ln(refined extended dictionary)
Panel A: ln(digital patent invention+1)							
Count	0.036***	0.038***	0.051***	0.038***	0.040***	0.049***	0.033***
	(0.006)	(0.007)	(0.009)	(0.007)	(0.008)	(0.009)	(0.006)
Observations	31,614	31,614	31,614	31,614	31,614	31,614	31,614
R-squared	0.760	0.760	0.760	0.760	0.760	0.760	0.760
Panel B: ln(digital patent+1)							
Count	0.034***	0.038***	0.044***	0.035***	0.040***	0.045***	0.032***
	(0.007)	(0.008)	(0.010)	(0.008)	(0.009)	(0.010)	(0.007)
Observations	31,614	31,614	31,614	31,614	31,614	31,614	31,614
R-squared	0.775	0.775	0.775	0.775	0.775	0.775	0.775
Controls & Constant	YES	YES	YES	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES

Note: This table shows the relationship between digital adoption intensity measures and digital patents. Dependent variable in Panel A is natural logarithm of the digital patent inventions plus one, while Panel B includes all the three types of digital patents. The columns represent different independent variables of digital adoption intensity: (1) ln(raw count) — the natural logarithm of the raw count of digital adoption mentions in the full annual report based on the primary digital-related keywords dictionaries, (2) ln(refined count) — the natural logarithm of the refined count after applying the refinement algorithm, (3) ln(digital scope) — the natural logarithm of the scope of digital technology adoption, (4) ln(raw count in MD&A) — the natural logarithm of raw digital adoption count specifically in the MD&A section, (5) ln(refined count in MD&A) — the refined count in the MD&A section, (6) ln(scope in MD&A) — the scope of digital adoption in the MD&A section, and (7) ln(refined extended dictionary) — the natural logarithm of digital adoption count in the MD&A section using the BERT-extended dictionary.

Table 6: Validity Checks: Software Assets and R&D Expenditure Ratio

	(1) ln(raw count)	(2) ln(refined count)	(3) ln(digital scope)	(4) ln(raw count in MD&A)	(5) ln(refined count in MD&A)	(6) ln(scope in MD&A)	(7) ln(refined extended dictionary)
Panel A: ln(software assets+1)							
Count	0.050***	0.056***	0.061**	0.039**	0.053***	0.045**	0.033***
	(0.019)	(0.017)	(0.024)	(0.016)	(0.016)	(0.019)	(0.006)
Observations	12,672	12,672	12,672	12,672	12,672	12,672	12,672
R-squared	0.899	0.899	0.899	0.899	0.899	0.899	0.899
Panel B: R&D expenditure/revenue							
Count	0.640*	0.704*	0.740	0.889**	1.020**	0.970*	0.806**
	(0.345)	(0.382)	(0.507)	(0.401)	(0.483)	(0.499)	(0.346)
Observations	20,411	20,411	20,411	20,411	20,411	20,411	20,411
R-squared	0.994	0.994	0.994	0.994	0.994	0.994	0.994
Controls & Constant	YES	YES	YES	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES

Note: This table shows the correlation between digital adoption intensity measures and software (Panel A) and R&D (Panel B). The dependent variable in Panel A is the natural logarithm of software assets plus one, while the dependent variable in Panel B is the rate of R&D expenditure over revenue. The independent variable in each column is the same as it in Table 5.

Table 7: Validity Checks: Intangible Assets Amortization

Dep. Var.: ln(intangible amortization+1)	(1) ln(raw count)	(2) ln(refined count)	(3) ln(digital scope)	(4) ln(raw count in MD&A)	(5) ln(refined count in MD&A)	(6) ln(scope in MD&A)	(7) ln(refined extended dictionary)
Count	0.030**	0.036***	0.022	0.039***	0.041***	0.036***	0.033***
	(0.012)	(0.011)	(0.016)	(0.010)	(0.011)	(0.014)	(0.009)
ln(intangible+1)	0.592***	0.592***	0.592***	0.596***	0.596***	0.596***	0.597***
	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)
Observations	31,614	31,614	31,614	31,614	31,614	31,614	31,614
R-squared	0.879	0.879	0.878	0.878	0.878	0.878	0.878
Controls & Constant	YES	YES	YES	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES

Note: This table shows the correlation between digital adoption intensity measures and amortization of intangible assets. The dependent variable is the natural logarithm of the amortization of intangible assets plus one. The independent variable in each column is the same as it in Table 5.

Table 8: Treatment Group Definition: 5-year and 6-year Horizons

Panel A: 5-Year Horizon

	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
Batch 2011	✗	✗	✓	✓	✓									
Batch 2012		✗	✗	✓	✓	✓								
Batch 2013			✗	✗	✓	✓	✓							
Batch 2014				✗	✗	✓	✓	✓						
Batch 2015					✗	✗	✓	✓	✓					
Batch 2016						✗	✗	✓	✓	✓				
Batch 2017							✗	✗	✓	✓	✓			
Batch 2018								✗	✗	✓	✓	✓		
Batch 2019									✗	✗	✓	✓	✓	
Batch 2020										✗	✗	✓	✓	✓

Panel B: 6-Year Horizon for robustness

	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
Batch 2012	✗	✗	✗	✓	✓	✓								
Batch 2013		✗	✗	✗	✓	✓	✓							
Batch 2014			✗	✗	✗	✓	✓	✓						
Batch 2015				✗	✗	✗	✓	✓	✓					
Batch 2016					✗	✗	✗	✓	✓	✓				
Batch 2017						✗	✗	✗	✓	✓	✓			
Batch 2018							✗	✗	✗	✓	✓	✓		
Batch 2019								✗	✗	✗	✓	✓	✓	
Batch 2020									✗	✗	✗	✓	✓	✓

Note: This table presents the definition of treatment group based on a 5-year horizon in Panel A as baseline setup and a 6-year horizon in Panel B later for robustness checks. Let T_{it} be a dummy variable that equals one if the refined occurrence of digital keywords for firm i in year t is greater than zero. For each batch, if the $T_{it} = 1$, it is indicated by a checkmark (✓); if $T_{it} = 0$, it is indicated by a cross (✗).

Table 9: Summary Statistics

Variable	Timing	Definition	N	mean	sd	p1	p5	p25	p50	p75	p95	p99
treatment		Dummy=1 if adopt in years t, $t+1$, and $t+2$, otherwise 0.	5020	0.153	0.36	0	0	0	0	0	1	1
batch	Year t	Batch	5020	2016	2.626	2012	2012	2014	2016	2018	2020	2020
age		Years since foundation	5020	17.75	5.405	5.5	8.5	14.5	17.5	21.5	26.5	30.5
size		Natural logarithm of no. employees	5020	7.572	1.248	4.605	5.659	6.729	7.502	8.313	9.91	10.73
cash ratio		(Cash and cash equivalents + financial assets held for trading) / total assets	5020	0.181	0.132	0.019	0.0398	0.0896	0.142	0.233	0.455	0.636
leverage		Total long-term liabilities / total equity	5020	0.433	0.215	0.0641	0.103	0.252	0.422	0.6	0.798	0.912
cip	Average	Construction in progress / total asset	5020	0.0491	0.0509	0.000172	0.00143	0.0146	0.0331	0.0643	0.158	0.247
sales growth	value of	Growth rate of sales	5020	0.138	0.183	-0.237	-0.112	0.031	0.113	0.218	0.459	0.8
sales growth sd	years	Standard deviation of three-year growth rates of sales	5020	0.243	0.406	0.0228	0.0418	0.0931	0.16	0.264	0.608	1.479
high skill	$t-2$ and $t-1$	No. employees with diploma and above / total no. employees	5020	0.442	0.209	0.0584	0.134	0.293	0.412	0.574	0.844	0.947
export		Dummy=1 if exported before digitization, otherwise 0	5020	0.474	0.499	0	0	0	0	1	1	1
SOE		Dummy=1 if State-Owned Enterprise, otherwise 0	5020	0.149	0.333	0	0	0	0	0	1	1
concentration		NAICS 2-digit industry-level concentration: share of sales, value of shipments, or revenue of top 50 largest firms in industry (US Census Bureau)	5020	0.582	0.194	0.142	0.154	0.474	0.637	0.709	0.808	0.954
industry adoption		Industry-level share of digital adoption	5020	0.219	0.137	0.0167	0.0228	0.118	0.194	0.332	0.463	0.59
lnlength		Natural logarithm of length of MD&A part of annual report	5020	8.855	0.555	7.386	7.854	8.528	8.901	9.233	9.703	9.993
internet		City-level broadband internet penetration rate (Cai and Wu, 2025)	5020	1.251	1.154	0.214	0.318	0.634	1.007	1.348	3.712	6.795
CEO IT		Dummy=1 if CEO has IT background, otherwise 0	5020	0.0219	0.137	0	0	0	0	0	0	1
lnEBITDA		Natural logarithm of earnings before interest, taxes, depreciation, and amortization (EBITDA)	5020	19.87	1.389	17.13	17.87	18.89	19.72	20.74	22.39	23.62
lnsales		Natural logarithm of prime operating revenue	5020	21.73	1.389	18.88	19.75	20.78	21.53	22.57	24.52	25.25
market share		Share of prime operating revenue in industry	5020	0.539	1.195	0.00346	0.00912	0.038	0.125	0.424	2.59	7.705
lnoversea	Average	Natural logarithm of overseas revenue plus one	5020	11.88	9.36	0	0	0	17.42	19.92	22.17	22.97
oversea share	value of	Overseas revenue / sales	5020	0.128	0.195	0	0	0	0.0264	0.177	0.594	0.797
COGS growth	years t,	Growth rate of cost of goods sold	5020	0.186	0.291	-0.223	-0.097	0.0326	0.122	0.242	0.735	1.4
labor cost growth	$t+1$,	Growth rate of cash paid to employees and on behalf of employees	5020	0.191	0.286	-0.0945	-0.0279	0.0573	0.126	0.228	0.576	1.985
management cost growth	and $t+2$	Growth rate of management cost	5020	0.112	0.174	-0.199	-0.116	0.00147	0.0815	0.189	0.465	0.647
TFP		Total Factor Productivity (TFP) using GMM method	5020	5.623	0.778	4.132	4.518	5.063	5.526	6.078	7.154	7.748
labor productivity		Natural logarithm of (prime operating revenue / no. employees)	5020	13.99	0.831	12.34	12.78	13.41	13.9	14.5	15.51	16.31
Tobin's Q		((Total shares - domestically listed foreign b-shares) * current closing price of a-shares + domestically listed foreign b-shares * current closing price of b-shares * current exchange rate) / total assets	5020	2.274	1.489	0.868	0.96	1.3	1.832	2.709	5.188	8.329
profitability		EBITDA / prime operating revenue	5020	0.204	0.157	0.028	0.0521	0.103	0.163	0.252	0.51	0.855
lnintangible		Natural logarithm of intangible assets	5020	18.78	1.879	12.41	15.55	17.87	18.8	19.8	21.98	22.64
lnamortization		Natural logarithm of amortization of long-term deferred expenses plus one	5020	11.79	5.719	0	0	9.891	14.01	15.53	17.72	18.69
employment growth		Growth rate of no. employees	5020	0.0808	0.164	-0.166	-0.0921	-0.0212	0.0346	0.132	0.448	0.635

Table 10: Probit Model: Key Determinants of Digital Technology Adoption

Dep. Var.: Adoption Dummy	(1) Coefficient	(2) Marginal Effect	(3) Coefficient	(4) Marginal Effect
age_{-t}	0.001	0.000	0.004	0.001
	(0.005)	(0.001)	(0.005)	(0.001)
$size_{-t}$	0.178***	0.034***	0.175***	0.035***
	(0.025)	(0.005)	(0.022)	(0.004)
cash $ratio_{-t}$	0.708***	0.136***	0.624***	0.126***
	(0.226)	(0.043)	(0.202)	(0.041)
$leverage_{-t}$	-0.159	-0.031	-0.047	-0.010
	(0.157)	(0.030)	(0.136)	(0.028)
construction in $progress_{-t}$	-1.313**	-0.253**	-1.345***	-0.272***
	(0.565)	(0.109)	(0.514)	(0.104)
sales $growth_{-t}$	0.528***	0.102***	0.485***	0.098***
	(0.154)	(0.030)	(0.141)	(0.029)
sales growth $volatility_{-t}$	-0.103	-0.020	-0.105	-0.021
	(0.083)	(0.016)	(0.076)	(0.015)
high skill $share_{-t}$	0.554***	0.107***	0.554***	0.112***
	(0.142)	(0.027)	(0.121)	(0.025)
$export_{-t}$	0.157***	0.030***	0.078	0.016
	(0.055)	(0.011)	(0.049)	(0.010)
SOE_{-t}	-0.230***	-0.044***	-0.213***	-0.043***
	(0.083)	(0.016)	(0.075)	(0.015)
$concentration_{-t}$			-0.387***	-0.078***
			(0.121)	(0.024)
industry $adoption_{-t}$			2.472***	0.501***
			(0.213)	(0.042)
MD&A $length_{-t}$	0.492***	0.095***	0.428***	0.087***
	(0.067)	(0.013)	(0.061)	(0.012)
internet $penetration_{-t}$	0.059*	0.011*	0.050*	0.010*
	(0.031)	(0.006)	(0.029)	(0.006)
CEO IT_{-t}	0.408***	0.079***	0.426***	0.086***
	(0.155)	(0.030)	(0.145)	(0.029)
Constant	-7.044***		-6.318***	
	(0.680)		(0.559)	
Province × Year FE	YES	YES	YES	YES
Industry × Year FE	YES	YES	NO	NO
Observations	5,020	5,020	5,211	5,211

Note: This table presents the Probit model results under 5-year horizon, with 2 years prior to the digital adoption, the adoption year, and 2 years after the adoption. Under the 5-year horizon, the treatment group for batch t is defined as firms with no digital adoption in years $t-2$ and $t-1$ and with digital adoption in years t, $t+1$, and $t+2$, while control group contains firms with no digital adoption from year $t-2$ to $t+2$. And subscript $-t$ denotes the average of year $t-2$ and $t-1$.

Table 11: Probit Model under Different Time Horizons

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.: Adoption Dummy	6-Year: 3+3		7-Year: 3+4		8-Year: 3+5	
age_{-t}	0.000	0.006	0.004	0.007	0.007	0.012
	(0.006)	(0.006)	(0.008)	(0.007)	(0.012)	(0.010)
$size_{-t}$	0.162***	0.153***	0.198***	0.184***	0.245***	0.187***
	(0.028)	(0.025)	(0.037)	(0.031)	(0.054)	(0.043)
cash $ratio_{-t}$	0.861***	0.692***	0.814**	0.840***	1.230**	1.219***
	(0.275)	(0.240)	(0.352)	(0.298)	(0.500)	(0.412)
$leverage_{-t}$	-0.128	0.037	-0.245	-0.067	0.033	0.289
	(0.179)	(0.153)	(0.235)	(0.193)	(0.332)	(0.264)
construction in $progress_{-t}$	-1.532**	-1.624***	-1.797**	-1.799**	-0.971	-1.667*
	(0.681)	(0.614)	(0.887)	(0.773)	(1.230)	(1.013)
sales $growth_{-t}$	0.581***	0.564***	0.703***	0.746***	1.133***	1.104***
	(0.195)	(0.177)	(0.253)	(0.221)	(0.375)	(0.316)
sales growth $volatility_{-t}$	-0.188	-0.175	-0.079	-0.060	-0.354	-0.524**
	(0.128)	(0.112)	(0.162)	(0.138)	(0.295)	(0.258)
high skill $share_{-t}$	0.429***	0.478***	0.777***	0.623***	0.791***	0.679***
	(0.164)	(0.138)	(0.215)	(0.177)	(0.301)	(0.242)
$export_{-t}$	0.246***	0.137**	0.284***	0.210***	0.409***	0.318***
	(0.063)	(0.054)	(0.081)	(0.069)	(0.116)	(0.094)
SOE_{-t}	-0.268***	-0.268***	-0.151	-0.174*	-0.263	-0.210
	(0.097)	(0.087)	(0.121)	(0.104)	(0.177)	(0.145)
$concerntration_{-t}$		-0.523***		-0.729***		-0.752***
		(0.136)		(0.171)		(0.231)
industry $adoption_{-t}$		2.193***		2.470***		2.271***
		(0.258)		(0.334)		(0.473)
MD&A $length_{-t}$	0.553***	0.473***	0.645***	0.521***	0.756***	0.584***
	(0.080)	(0.072)	(0.105)	(0.092)	(0.151)	(0.126)
internet $penetration_{-t}$	0.084**	0.075**	0.103**	0.094**	0.099	0.086
	(0.036)	(0.034)	(0.045)	(0.041)	(0.061)	(0.055)
CEO IT_{-t}	0.601***	0.560***	1.066***	0.915***	1.149***	1.102***
	(0.187)	(0.171)	(0.244)	(0.222)	(0.377)	(0.338)
Constant	-7.150***	-6.326***	-8.775***	-7.241***	-10.273***	-8.315***
	(0.778)	(0.644)	(1.020)	(0.827)	(1.424)	(1.146)
Province × Year FE	YES	YES	YES	YES	YES	YES
Industry × Year FE	YES	NO	YES	NO	YES	NO
Observations	4,120	4,363	2,710	2,949	1,460	1,673

Note: This table is the Probit model result under 6-year, 7-year, and 8-year horizons. Columns (1) and (2) use a 6-year horizon, including 3 years before, the adoption year, and 2 years after adoption. Columns (3) and (4) extend the post-adoption period, covering 3 years before, the adoption year, and 3 years after adoption.Columns (5) and (6) focus on the 4-year period following adoption. Under the 6-year horizon, the treatment group for batch t is defined as firms with no digital adoption in years $t-3$, $t-2$ and $t-1$ and with digital adoption in years t, $t+1$, and $t+2$, while control group contains firms with no digital adoption from year $t-3$ to $t+2$. And subscript $-t$ denotes the average of year $t-3$, $t-2$ and $t-1$. Similar rule applies to 7-year and 8-year horizons.

Table 12: Outcome Model: Firm Profit after Digital Technology Adoption

	(1)	(2)	(3)	(4)	(5)	(6)
		MLE			CF	
	Simple Model	General Model		Simple Model	General Model	
Dep. Var.: ln(EBITDA)	All	Control	Treatment	All	Control	Treatment
treatment	0.427**			0.463***		
	(0.171)			(0.173)		
age	0.004*	0.005**	-0.001	0.004*	0.006**	-0.005
	(0.002)	(0.002)	(0.005)	(0.002)	(0.002)	(0.008)
size	0.875***	0.873***	0.867***	0.874***	0.870***	0.881***
	(0.013)	(0.013)	(0.042)	(0.013)	(0.014)	(0.046)
cash ratio	0.048	0.094	-0.097	0.036	0.092	-0.280
	(0.124)	(0.136)	(0.320)	(0.128)	(0.142)	(0.419)
leverage	0.070	0.046	0.257	0.080	0.067	0.368
	(0.072)	(0.077)	(0.177)	(0.074)	(0.079)	(0.304)
construction in progress	0.926***	0.956***	0.606	0.932***	0.903***	1.697*
	(0.242)	(0.255)	(0.784)	(0.251)	(0.269)	(1.019)
sales growth	0.967***	1.080***	0.429***	0.974***	1.091***	0.339
	(0.077)	(0.084)	(0.159)	(0.079)	(0.088)	(0.221)
sales growth volatility	-0.118***	-0.115***	-0.110	-0.117***	-0.118***	-0.039
	(0.037)	(0.040)	(0.093)	(0.038)	(0.043)	(0.112)
high skill share	1.437***	1.410***	1.464***	1.434***	1.398***	1.227***
	(0.066)	(0.070)	(0.162)	(0.067)	(0.072)	(0.235)
export	-0.017	-0.026	0.017	-0.019	-0.028	0.042
	(0.023)	(0.024)	(0.065)	(0.023)	(0.025)	(0.081)
SOE	0.058	0.046	0.162	0.061	0.056	0.188
	(0.040)	(0.044)	(0.105)	(0.042)	(0.046)	(0.147)
constant	12.929***	12.886***	13.439***	12.923***	12.888***	13.306***
	(0.184)	(0.190)	(0.769)	(0.185)	(0.199)	(0.812)
ρ	-0.361***	-0.545***	-0.251			
	(0.123)	(0.107)	(0.384)			
σ	0.708***	0.717***	0.729***			
	(0.014)	(0.016)	(0.058)			
β_2				-0.498***	-0.726***	-0.205
				(0.175)	(0.211)	(0.459)
Province × Year FE	YES	YES	YES	YES	YES	YES
Industry × Year FE	YES	YES	YES	YES	YES	YES
Observations	5,020	4,254	766	5,020	4,254	766

Note: This table is the outcome model result under the 5-year horizon, with 2 years prior to the digital adoption, the adoption year, and 2 years after the adoption. Under the 5-year horizon, the treatment group for batch t is defined as firms with no digital adoption in years $t-2$ and $t-1$ and with digital adoption in years t, $t+1$, and $t+2$, while control group contains firms with no digital adoption from years $t-2$ to $t+2$. The dependent variable is the average value of the natural logarithm of EBITDA in years t, $t+1$, and $t+2$. Columns (1) to (3) are estimated by MLE method, while Columns (4) to (6) are estimated by CF approach. At the bottom of the Columns (1) to (3), ρ is the estimated correlation between the treatment-assignment error (v_i) and the outcome error (ε_i), and σ is the standard error of ε_i. In Columns (4) to (6), β_2 is in Equation 40, which shows the direction of selection bias.

Table 13: Decomposition: Treatment Effects and Selection Bias

Quantity of Interest	(1) Simple model	(2) General model
$E[y_{i0}\|A_i=0]$ - observed	19.824	19.824
$E[y_{i1}\|A_i=1]$ - observed	20.124	20.124
$E[y_{i1}\|A_i=0]$ - predicted	20.287	20.005
$E[y_{i0}\|A_i=1]$ - predicted	19.661	19.454
$E[y_{i0}]$ - potential outcome mean	19.799	19.768
$E[y_{i1}]$ - potential outcome mean	20.262	20.023
$ATE = E[y_{i1}] - E[y_{i0}]$	0.463***	0.255
$ATE/E[y_{i0}]$	2.34%	1.29%
$ATET = E[y_{i1}\|A_i=1] - E[y_{i0}\|A_i=1]$	0.463***	0.670***
$\boldsymbol{ATET/E[y_{i1}\|A_i=1]}$	**2.30%**	**3.33%**
$ATNT = E[y_{i1}\|A_i=0] - E[y_{i0}\|A_i=0]$	0.463***	0.181
$ATNT/E[y_{i0}\|A_i=0]$	2.33%	0.91%
$GMD = E[y_{i1}\|A_i=1] - E[y_{i0}\|A_i=0]$	0.300***	0.300***
$ATET = E[y_{i1}\|A_i=1] - E[y_{i0}\|A_i=1]$	0.463***	0.670***
obs: $(\beta_{11} - \beta_{10}) \times E[x_i\|A_i=1]$	0.463***	0.305
unobs: $(\beta_{21} - \beta_{20}) * E[v_i\|A_i=1]$	0	0.365
$SB = E[y_{i0}\|A_i=1] - E[y_{i0}\|A_i=0]$	-0.163***	-0.370*
obs: $\beta_{10} \times (E[x_i\|A_i=1] - E[x_i\|A_i=0])$	0.249***	0.231***
unobs: $\beta_{20} \times (E[v_i\|A_i=1] - E[v_i\|A_i=0])$	-0.412***	-0.601***
$S_0 = E[y_{i0}\|A_i=1] - E[y_{i0}]$	-0.138***	-0.314*
$S_1 = E[y_{i1}\|A_i=1] - E[y_{i1}]$	-0.138***	0.101

Note: This table presents the decomposition of the treatment effect and selection bias of digital adoption on firm profit bsed on the control function estimation results in Columns (5) and (6) of Table 12. The dependent variable, y_i, represents the natural logarithm of EBITDA, where the subscript 0 indicates the non-adopting status, and 1 represents the adopting status. Significance levels are obtained from bootstrap standard errors.

Table 14: Time-Varying Effects of Digital Technology Adoption on Firm Profit

Dep. Var.: ln(EBITDA)	(1) MLE	(2) CF
Panel A: Adoption year t		
ATE	0.472***	0.612***
	(0.149)	(0.180)
Panel B: Adoption year $t+1$		
ATE	0.509***	0.671***
	(0.128)	(0.172)
Panel C: Adoption year $t+2$		
ATE	0.383***	0.500***
	(0.143)	(0.173)
Three-year average $E[y_{i0}]$	19.805	19.768
Observations	5,020	5,020
Controls and Constant	YES	YES
Province × Year FE	YES	YES
Industry × Year FE	YES	YES

Note: This table shows the time-varying effects of digital adoption on profit under the 5-year horizon. The dependent variables in Panels A, B, and C are the natural logarithm of EBITDA in years t, $t+1$, and $t+2$, respectively. The treatment group for batch t is defined as firms with no digital adoption in years $t-2$ and $t-1$ and with digital adoption in years t, $t+1$, and $t+2$, while control group contains firms with no digital adoption from years $t-2$ to $t+2$. Column (1) is estimated by MLE, while Column (2) is estimated by CF approach. Three-year average $E[y_{i0}]$ refers to the three-year average value of the natural logarithm of predicted non-adopting EBITDA.

Table 15: Mechanism: Demand Expansion

Dep. Var.	(1) ln(Sales)	(2) Market Share	(3) ln(Overseas)	(4) Overseas Share
Panel A: Adoption year t				
ATE	0.543***	1.023***	3.509**	0.089**
	(0.148)	(0.204)	(1.575)	(0.041)
R-squared	0.793	0.486	0.490	0.208
Panel B: Adoption year $t+1$				
ATE	0.738***	1.069***	4.406***	0.112***
	(0.144)	(0.197)	(1.538)	(0.040)
R-squared	0.792	0.481	0.495	0.209
Panel C: Adoption year $t+2$				
ATE	0.778***	1.151***	5.185***	0.130***
	(0.141)	(0.193)	(1.520)	(0.039)
R-squared	0.797	0.480	0.498	0.213
Three-year average $E[y_{i0}]$	21.649	0.393	11.232	0.111
Observations	5,020	5,020	5,020	5,020
Controls and Constant	YES	YES	YES	YES
Province × Year FE	YES	YES	YES	YES
Industry × Year FE	YES	YES	YES	YES

Note: This table presents the CF estimation results for the mechanism of demand expansion over a 5-year horizon. The layout follows the structure of Table 15, with the dependent variables listed at the top of each column. Each cell within a column and panel corresponds to an individual regression. Specifically, in Column (1) of Panels A, B, and C, the dependent variables are the natural logarithm of operating revenue for years t, $t+1$, and $t+2$, respectively. In Column (2) of Panels A, B, and C, the dependent variables represent the market share, defined as the share of operating revenue within the industry for years t, $t+1$, and $t+2$, respectively. Similarly, in Column (3) of Panels A, B, and C, the dependent variables are the natural logarithm of overseas revenues plus one for years t, $t+1$, and $t+2$, respectively. Finally, in Column (4) of Panels A, B, and C, the dependent variables represent the ratio of overseas revenues to total revenues for years t, $t+1$, and $t+2$, respectively. Three-year average $E[y_{i0}]$ in each column refers to the average value of the dependent variables in year t, $t+1$, and $t+2$ estimated by the CF approach.

Table 16: Mechanism: Cost Reduction

Dep. Var.	(1) Δln(COGS)	(2) Δln(Labor Cost)	(3) Δln(Management Cost)
Panel A: Adoption year t			
ATE	-0.468***	-0.965***	-0.185***
	(0.090)	(0.101)	(0.058)
R-squared	0.384	0.300	0.322
Panel B: Adoption year $t+1$			
ATE	-0.132	-0.308***	0.021
	(0.085)	(0.095)	(0.054)
R-squared	0.352	0.235	0.297
Panel C: Adoption year $t+2$			
ATE	-0.044	-0.263***	0.045
	(0.087)	(0.077)	(0.053)
R-squared	0.355	0.266	0.282
Three-year average $E[y_{i0}]$	0.213	0.250	0.110
Observations	5,020	5,020	5,020
Controls and Constant	YES	YES	YES
Province × Year FE	YES	YES	YES
Industry × Year FE	YES	YES	YES

Note: This table presents the CF estimation results for the mechanism of cost reduction over a 5-year horizon. The layout mirrors that of Table 15, with the dependent variables specified at the top of each column. Specifically, in Column (1) of Panels A, B, and C, the dependent variables are the growth of COGS for years t, $t+1$, and $t+2$, respectively. In Column (2) of Panels A, B, and C, the dependent variables are the growth of labor cost for years t, $t+1$, and $t+2$, respectively. Finally, in Column (3) of Panels A, B, and C, the dependent variables are the growth of management cost for years t, $t+1$, and $t+2$, respectively.

Table 17: Mechanism: Demand Expansion and Cost Reduction in Demand-Driven and Production-Driven Industries

Dep. Var.	(1)	(2)	(3)	(4)	(5)	(6)
		ln(Sales)			Δln(COGS)	
	All Industries	Demand Driven	Production Driven	All Industries	Demand Driven	Production Driven
Panel A: Adoption year t						
ATE	0.543***	0.684**	0.464***	-0.468***	-0.336	-0.302**
	(0.148)	(0.284)	(0.158)	(0.090)	(0.207)	(0.121)
R-squared	0.793	0.783	0.810	0.384	0.373	0.427
Panel B: Adoption year $t+1$						
ATE	0.738***	0.924***	0.514***	-0.132	-0.139	-0.200**
	(0.144)	(0.275)	(0.153)	(0.085)	(0.163)	(0.090)
R-squared	0.792	0.782	0.810	0.352	0.473	0.342
Panel C: Adoption year $t+2$						
ATE	0.778***	0.970***	0.437***	-0.044	0.317	-0.164**
	(0.141)	(0.279)	(0.150)	(0.087)	(0.216)	(0.081)
R-squared	0.797	0.782	0.814	0.355	0.405	0.355
Three-year average $E[y_{i0}]$	21.649	21.558	21.670	0.213	0.210	0.210
Observations	5,020	1,403	3,098	5,020	1,403	3,098
Controls and Constant	YES	YES	YES	YES	YES	YES
Province × Year FE	YES	YES	YES	YES	YES	YES
Industry × Year FE	YES	YES	YES	YES	YES	YES

Note: This table presents the subgroup CF estimation results for demand-driven and production-driven industries. The dependent variables in Columns (1) to (3) of Panels A, B, and C are the natural logarithm of operating revenue in years t, $t+1$, and $t+2$, respectively. In Columns (4) to (6) of Panels A, B, and C, the dependent variables are the growth of COGS in years t, $t+1$, and $t+2$, respectively. Columns (1) and (4) use the full sample as the baseline. Columns (3) and (6) focus on firms in production-driven industries, identified by WIND 4-digit industry classification codes 1010 (Energy), 1510 (Materials), 2010 (Capital Goods), 2510 (Automobiles and Components), 3510 (Health Care Equipment and Services), and 3520 (Pharmaceuticals, Biotechnology, and Life Sciences). Columns (2) and (5) focus on firms in demand-driven industries, encompassing all remaining industries. The sum of observations in Columns (2) and (3) does not equal that of Column (1) because, in the Probit model with rich fixed-effects (FE) interactions, some FE groups exhibit no variation in the adoption dummy. For instance, in province 62 in 2012, all four observations had a zero adoption dummy, resulting in the exclusion of this FE group and the omission of these observations. As a result, the more granular the subgroup, the more observations are omitted, leading to a smaller total sample size.

Table 18: Mechanism: Productivity Enhancement

Dep. Var.	(1) TFP (OP)	(2) TFP (LP)	(3) TFP (GMM)	(4) Labor Prod
Panel A: Adoption year t				
ATE	0.475***	0.536***	0.569***	0.457***
	(0.138)	(0.138)	(0.132)	(0.145)
R-squared	0.515	0.654	0.491	0.462
Panel B: Adoption year $t+1$				
ATE	0.624***	0.693***	0.731***	0.633***
	(0.134)	(0.134)	(0.128)	(0.142)
R-squared	0.520	0.658	0.497	0.466
Panel C: Adoption year $t+2$				
ATE	0.633***	0.699***	0.723***	0.683***
	(0.130)	(0.132)	(0.125)	(0.138)
R-squared	0.525	0.659	0.499	0.469
Three-year average $E[y_{i0}]$	6.990	8.705	5.541	13.925
Observations	5,020	5,020	5,020	5,020
Controls and Constant	YES	YES	YES	YES
Province × Year FE	YES	YES	YES	YES
Industry × Year FE	YES	YES	YES	YES

Note: This table presents the CF estimation results for the mechanism of productivity enhancement over a 5-year horizon. The layout mirrors that of Table 15. The dependent vairbles in Columns (1) to (3) are TFPs estimated by Olley and Pakes (1992) , Levinsohn and Petrin (2003), and Blundell and Bond (2000), respectively. And the labour productivity in Column (4) is measured by the natural logarithm of prime operating revenue over the employees.

Table 19: Mechanism: Capital Labor Substitution

Dep. Var.	(1) ln(Intangible)	(2) ln(Amortization)	(3) Δln(Employment)
Panel A: Adoption year t			
ATE	1.173***	5.369***	-0.450***
	(0.311)	(1.273)	(0.062)
R-squared	0.535	0.233	0.200
Panel B: Adoption year $t+1$			
ATE	1.303***	5.133***	-0.136**
	(0.295)	(1.205)	(0.054)
R-squared	0.545	0.229	0.176
Panel C: Adoption year $t+2$			
ATE	1.122***	5.609***	-0.066
	(0.283)	(1.151)	(0.050)
R-squared	0.549	0.215	0.191
Three-year average $E[y_{i0}]$	18.580	10.956	0.099
Observations	5,020	5,020	5,020
Controls and Constant	YES	YES	YES
Province × Year FE	YES	YES	YES
Industry × Year FE	YES	YES	YES

Note: This table presents the CF estimation results for the capital-labor substitution over a 5-year horizon. The layout mirrors that of Table 15, with the dependent variables specified at the top of each column.

Table 20: Mechanism: Market Valuation and Firm Profitability

Dep. Var.	(1) Tobin's Q	(2) Profitability
Panel A: Adoption year t		
ATE	0.690**	-0.051
	(0.292)	(0.031)
R-squared	0.423	0.340
Panel B: Adoption year $t+1$		
ATE	0.476	-0.047
	(0.293)	(0.031)
R-squared	0.397	0.346
Panel C: Adoption year $t+2$		
ATE	0.380	-0.084**
	(0.290)	(0.034)
R-squared	0.372	0.349
Three-year average $E[y_{i0}]$	2.232	0.211
Observations	5,020	5,020
Controls and Constant	YES	YES
Province × Year FE	YES	YES
Industry × Year FE	YES	YES

Note: This table presents the CF estimation results for market valuation and firm profitability over a 5-year horizon. The layout mirrors that of Table 15.

Reference

Acemoglu, D., Anderson, G., Beede, D., Buffington, C., Childress, E., Dinlersoz, E., Foster, L., Goldschlag, N., Haltiwanger, J., Kroff, Z., Restrepo, P. and Zolas, N. (2023), 'Advanced technology adoption: Selection or causal effects?', *American Economic Association: Papers and Proceedings* **113**, 210–214.

Acemoglu, D., Anderson, G. W., Beede, D. N., Buffington, C., Childress, E. E., Dinlersoz, E., Foster, L. S., Goldschlag, N., Haltiwanger, J. C., Kroff, Z. et al. (2022), 'Automation and the workforce: A firm-level view from the 2019 annual business survey', *National Bureau of Economic Research* .

Acemoglu, D. and Restrepo, P. (2020), 'Robots and jobs: Evidence from us labor markets', *Journal of Political Economy* **128**(6), 2188–2244.

Anderton, R., Jarvis, V., Labhard, V., Petroulakis, F. and Vivian, L. (2020), 'Virtually everywhere? digitalisation and the euro area and eu economies', *ECB Occasional paper* (2020244).

Babina, T., Fedyk, A., He, A. and Hodson, J. (2024), 'Artificial intelligence, firm growth, and product innovation', *Journal of Financial Economics* **151**, 103745.

Ban, G.-Y. and Keskin, N. B. (2021), 'Personalized dynamic pricing with machine learning: High-dimensional features and heterogeneous elasticity', *Management Science* **67**(9), 5549–5568.

Becchetti, L., Bedoya, D. A. L. and Paganetto, L. (2003), 'Ict investment, productivity and efficiency: evidence at firm level using a stochastic frontier approach', *Journal of productivity analysis* **20**, 143–167.

Bessen, J. (2018), Ai and jobs: The role of demand, Technical report, National Bureau of Economic Research.

Blundell, R. and Bond, S. (2000), 'Gmm estimation with persistent panel data: an application to production functions', *Econometric reviews* **19**(3), 321–340.

Borjas, G. J. (1987), Self-selection and the earnings of immigrants, Technical report, National Bureau of Economic Research.

Borjas, G. J. (1991), Immigration and self-selection, *in* 'Immigration, trade, and the labor market', University of Chicago Press, pp. 29–76.

Bresnahan, T. F. (1989), 'Empirical studies of industries with market power', *Handbook of industrial organization* **2**, 1011–1057.

Chen, W. and Srinivasan, S. (2023), 'Going digital: Implications for firm value and performance', *Review of Accounting Studies* pp. 1–47.

Cong, L. W., Liang, T. and Zhang, X. (2019), 'Textual factors: A scalable, interpretable, and data-driven approach to analyzing unstructured information', *Interpretable, and Data-driven Approach to Analyzing Unstructured Information (September 1, 2019)* .

Cong, L. W., Yang, X. and Zhang, X. (2024), 'Small and medium enterprises amidst the pandemic and reopening: Digital edge and transformation', *Management Science* .

Dubé, J.-P. and Misra, S. (2023), 'Personalized pricing and consumer welfare', *Journal of Political Economy* **131**(1), 131–189.

Goldfarb, A. and Tucker, C. (2019), 'Digital economics', *Journal of Economic Literature* **57**(1), 3–43.

Granados, N., Gupta, A. and Kauffman, R. J. (2012), 'Online and offline demand and price elasticities: Evidence from the air travel industry', *Information Systems Research* **23**(1), 164–181.

Greenwood, J., Ma, Y. and Yorukoglu, M. (2024), '"you will:" a macroeconomic analysis of digital advertising', *Review of Economic Studies* p. rdae067.

Griffin, H. F. and Dugan, M. T. (2003), 'Systematic risk and revenue volatility', *Journal of Financial Research* **26**(2), 179–189.

Gu, Z., Tang, S. and Wu, D. (2020), 'The political economy of labor employment decisions: Evidence from china', *Management Science* **66**(10), 4703–4725.

Haislip, J. Z. and Richardson, V. J. (2018), 'The effect of ceo it expertise on the information environment: Evidence from earnings forecasts and announcements', *Journal of Information Systems* **32**(2), 71–94.

Heckman, J. J. (1979), 'Sample selection bias as a specification error', *Econometrica: Journal of the econometric society* pp. 153–161.

Hjort, J. and Poulsen, J. (2019), 'The arrival of fast internet and employment in africa', *American Economic Review* **109**(3), 1032–1079.

Hjort, J. and Tian, L. (2024), 'The economic impact of internet connectivity in developing countries', *INSEAD Working paper* .

Hu, Q., Li, W., Lin, C. and Wei, L. (2024), 'What causes privatization? evidence from import competition in china', *Management Science* **70**(5), 3080–3101.

King, G., Lam, P. and Roberts, M. E. (2017), 'Computer-assisted keyword and document set discovery from unstructured text', *American Journal of Political Science* **61**(4), 971–988.

Koch, M., Manuylov, I. and Smolka, M. (2021), 'Robots and firms', *The Economic Journal* **131**(638), 2553–2584.

Law, C. C. and Ngai, E. W. (2007), 'Erp systems adoption: An exploratory study of the organizational factors and impacts of erp success', *Information & Management* **44**(4), 418–432.

Levinsohn, J. and Petrin, A. (2003), 'Estimating production functions using inputs to control for unobservables', *The review of economic studies* **70**(2), 317–341.

Lin, J. Y., Cai, F. and Li, Z. (1998), 'Competition, policy burdens, and state-owned enterprise reform', *The American economic review* **88**(2), 422–427.

Olley, G. S. and Pakes, A. (1992), The dynamics of productivity in the telecommunications equipment industry, Working Paper 3977, National Bureau of Economic Research.

Sautner, Z., Van Lent, L., Vilkov, G. and Zhang, R. (2023), 'Firm-level climate change exposure', *The Journal of Finance* **78**(3), 1449–1498.

Weyl, E. G. and Fabinger, M. (2013), 'Pass-through as an economic tool: Principles of incidence under imperfect competition', *Journal of political economy* **121**(3), 528–583.

Wooldridge, J. M. (2010), *Econometric analysis of cross section and panel data*, MIT press.