*Economic Society and Humanities Vol. 3 No. 7, 2026*

Academic Education Publishing House

# The Unreliable Impact of Digital Transformation on Firm Performance: A Theoretical Framework and Empirical Test Based on Chinese Listed Firms

**Zijian Zhou**

*University of Birmingham, Birmingham, B15 2TT, United Kingdom*

**Abstract: Two decades of firm-level research have produced a strikingly unstable answer to a simple question: does digital transformation pay? Estimates of the effect of digital transformation on firm performance range from strongly positive to statistically indistinguishable from zero to significantly negative, often within the same national setting and the same time window. This paper argues that the instability is not primarily a data problem or an identification problem, but a theoretical one. Digital technologies are general purpose technologies whose returns are conditional on the co-invention of complementary intangible assets - reorganised workflows, redistributed decision rights, upgraded human capital, and reconstructed data governance. I formalise the distance between a firm's accumulated digital technology stock and its stock of complementary organisational capital as the *complementarity gap*, and argue that this latent quantity, rather than digital intensity itself, governs performance outcomes. The framework generates five testable propositions concerning the average effect, its temporal shape, its moderation by organisational capital and environmental turbulence, and its distributional consequences across the productivity frontier. The author then develops a full empirical research design to test these propositions on Chinese A-share listed firms over 2013–2024, combining text-based measurement of digital transformation with disclosed human-capital and organisational-structure data, and combining staggered difference-in-differences estimation around the Broadband China pilot programme with heterogeneity-robust estimators, shift-share instrumentation, and threshold regression. The paper contributes a reconciling framework for a fragmented literature, an operationalisable construct that can be measured with existing archival data, and a research protocol that other scholars can adopt or contest.**

**Keywords: Digital Transformation; Firm Performance; Complementary Assets; Staggered Difference-in-Differences; Chinese listed Firms**

## 1. Introduction

Few propositions in contemporary management practice command as much consensus as the claim that firms must digitalise or decline. That consensus is expensive. It is also, on the evidence, insufficiently examined [1].

The macroeconomic backdrop is not in dispute. In China, the value added of the core industries of the digital economy reached 14.09 trillion yuan in 2024, equivalent to 10.5 per cent of gross domestic product, up from 9.9 per cent a year earlier and comfortably ahead of the 10 per cent target set for 2025 under the 14th Five-Year Plan. Adoption of the newest layer of digital capability has been faster still: by late 2024, roughly two-fifths of the United States working-age population reported using generative artificial intelligence, and just under a quarter of employed respondents had used it for work in the preceding week — a diffusion rate at least as rapid as that of the personal computer measured from first mass-market launch. Firm-level rollouts have delivered measurable operational gains in well-defined settings [2,3]; access to an AI assistant raised issue resolution rates by roughly 15 per cent in one large customer-support deployment, with disproportionate benefits accruing to less experienced workers.

Yet at the level of the individual firm, the return on digital investment remains contested. The most recent meta-analytic synthesis of the field, pooling 74 primary studies, confirms a positive average association between digital transformation and performance outcomes but simultaneously documents that the magnitude and even the direction of reported effects vary systematically with how digital transformation is measured, which industry is studied, and which national economy supplies the sample. This is a diplomatic way of saying that the literature does not agree with itself. Some studies find that digital transformation raises operating performance and market valuation; others find null effects on the same outcomes; others still report negative effects on risk-taking and profitability [4]. The pattern is not new. An almost identical arc played out a generation earlier in the information technology business value literature, where an influential demonstration that superior IT capability was associated with superior financial performance was, fourteen years later and in the same journal, met with a careful replication finding no significant link in more recent data.

The standard response to contradictory findings is methodological: better instruments, cleaner shocks, richer controls. That response is necessary but, I will argue, not sufficient. The deeper difficulty is that the field has largely modelled digital transformation as an input whose marginal product can be estimated, when it is better understood as a *general purpose technology* whose marginal product is undefined in isolation. General purpose technologies do not raise output on their own; they raise output when paired with complementary investments that are typically intangible, firm-specific, slow to accumulate, and poorly captured in accounts. Because those complementary investments are expensed as incurred while their returns arrive with a lag, measured productivity first falls and only later rises — the Productivity J-curve [5,6]. Adjusting United States national accounts for intangibles related to computer hardware and software yields a total factor productivity level 15.9 per cent above official measures by the end of 2017, an indication of how much economic activity the standard measurement apparatus fails to see.

If this is right, then the question "does digital transformation improve firm performance?" is underspecified in the same way that "does adding fertiliser improve yield?" is underspecified without reference to soil, water, and seed. The answer depends on what else the firm has. This paper takes that intuition seriously and develops it into a construct that can be measured and a design that can test it [7].

I define the complementarity gap as the distance between a firm's accumulated digital technology stock and its stock of complementary organisational capital — the human capital, decision-rights architecture, process redesign, and data governance required to convert digital capability into economic value. My central claim is that the complementarity gap, not digital intensity, is the proximate determinant of whether digitalisation raises or depresses measured performance. Where the gap is wide, digital investment predominantly generates expensed cost with deferred and partly unmeasured benefit, and the estimated effect on contemporaneous performance is weak or negative [8]. Where the gap is narrow, the same digital investment is converted into measurable margin, productivity, and valuation gains. Because firms differ enormously in their stock of complementary assets, and because the standard text-based measure of digital transformation captures technology adoption and disclosure while remaining largely blind to complementary reorganisation, the field has been estimating an interaction effect using only one of its two arguments. Unstable coefficients are the predictable consequence.

This paper makes three contributions. Theoretically, it supplies a framework that reconciles the positive, null, and negative findings in the literature as different regions of a single underlying relationship rather than as competing claims requiring adjudication. It links the strategic management literature on dynamic capabilities and the economics literature on technological complementarity in a way that generates directional, falsifiable predictions rather than a general injunction to build capabilities [9-11]. Methodologically, it operationalises the complementarity gap using variables that are already disclosed in Chinese listed-firm filings and available in standard commercial databases, making the construct testable without proprietary survey access. Practically, it derives implications for sequencing — specifically, that the marginal return to digital investment is increasing in complementary organisational capital, which implies that firms below a capability threshold should invest in organisation before technology, and that policy instruments subsidising technology adoption without addressing complementary capability may widen rather than narrow the productivity distribution.

The remainder of the paper proceeds as follows. Section 2 reviews the conceptual and empirical literature and locates the gap. Section 3 develops the theoretical framework and states five propositions. Section 4 presents the empirical research design in detail, including measurement, identification, and robustness. Section 5 discusses anticipated findings and their theoretical, managerial, and policy implications. Section 6 concludes [12-14].

## 2. Literature Review

### 2.1 What Digital Transformation is, and What it is Not

Terminological inflation has been a persistent obstacle in this field. Practitioners and scholars alike have used "digital transformation" to describe activities ranging from replacing paper forms with electronic ones to reconstituting a firm's entire business model. A useful three-stage distinction separates *digitisation*, the conversion of analogue information into digital form; *digitalisation*, the use of digital technologies to alter existing processes; and *digital transformation* proper, in which the firm's business model and value logic are themselves reconstituted. The three stages differ not merely in degree but in the extent of complementary organisational change they require — a point that will do substantial work in Section 3.

The most widely adopted synthesis in the information systems literature, built inductively from a review of 282 works, characterises digital transformation as a process in which digital technologies generate disruptions that trigger strategic responses, whereby organisations alter their value creation paths while managing the structural changes and organisational barriers that condition whether the outcomes are positive or negative [15]. A parallel systematic review of 279 articles in the management literature reaches a compatible conclusion, finding that digital transformation pushes firms toward malleable organisational designs embedded in digital business ecosystems, and — importantly for present purposes — that conventional frameworks of organisational change capture the phenomenon only partially [16].

Two features of these definitions deserve emphasis. First, both treat digital transformation as a *process* rather than a state, which implies that cross-sectional comparisons of "transformed" and "untransformed" firms are conceptually awkward. Second, both explicitly place organisational barriers on the causal path between technology and outcome. The literature's own foundational definitions, in other words, already contain the conditionality that the empirical literature has largely failed to model.

### 2.2 Three Theoretical Lenses

2.2.1 The resource-based view

The resource-based view attributes performance differentials to resources that are valuable, rare, imperfectly imitable, and non-substitutable. Applied to technology, its central insight is that technology alone rarely satisfies these conditions — hardware and standard software are purchasable by competitors — whereas the organisational capability to deploy technology may. The canonical application classified firm-specific information technology resources into infrastructure, human IT resources, and IT-enabled intangibles, and found that firms with superior IT capability outperformed matched controls on a range of profit and cost measures. The framework's logic points directly toward complementarity: it is the bundle, not the technology, that is inimitable. [17]

2.2.2 Dynamic capabilities

Where the resource-based view explains why some bundles yield rents, the dynamic capabilities perspective explains how bundles are reconfigured under environmental change. The framework decomposes capability into sensing opportunities and threats, seizing them through investment and business model choice, and reconfiguring the asset base as conditions shift, and locates these capacities in microfoundations — skills, processes [18], decision rules, organisational structures — that are difficult to develop and deploy. Qualitative work applying this lens directly to digital transformation in incumbent firms identifies nine microfoundations and characterises digital transformation as an ongoing process of strategic renewal in which agility functions as the core mechanism linking digital technology to business model change. The theoretical implication is that the same technology yields different returns depending on the reconfiguration capacity of the adopting organisation.

2.2.3 Complementarity economics

The economics literature supplies the sharpest formal statement. Where activities are complements, the returns to increasing one are increasing in the level of the others, so that partial adoption of a system of mutually reinforcing practices can be worse than no adoption at all. Firm-level evidence bears this out for information technology specifically: complementarities appear among information technology, workplace reorganisation, and new product and service introduction, and the labour-demand and productivity effects of information technology are substantially larger when it is combined with the identified organisational investments. Related work quantifies the resulting intangible stock, showing that computer-related organisational capital is large and is capitalised by financial markets even though it does not appear on balance sheets. This literature also supplies the temporal prediction: because complementary intangible investment is expensed while its output is unmeasured [19], measured productivity growth is understated early in the diffusion of a general purpose technology and overstated later, generating the J-curve.

These three lenses are usually deployed separately. Their combination is more than additive. The resource-based view explains *why* complementarity generates rents; dynamic capabilities explain *how* complementary assets are built; complementarity economics explains *what shape* the resulting performance relationship should take over time and across firms. The framework in Section 3 uses all three.

### 2.3 Three Waves of Empirical Evidence

The empirical literature can be read as three waves, each rediscovering the same problem.

*Wave one: the productivity paradox.* Following Solow's much-quoted observation that the computer age was visible everywhere except in the productivity statistics, a substantial literature attempted to locate the returns to information technology investment. Firm-level work eventually recovered substantial returns that aggregate data had obscured, and an influential review synthesised the field into an integrative model of information technology business value in which organisational and environmental factors mediate the technology-performance relationship [20].

*Wave two: information technology capability.* Attention shifted from spending to capability. The finding that IT-capable firms outperformed matched controls was widely replicated and became a settled result — until a careful reassessment using data from the 2000s found no significant link, and attributed the divergence to the commoditisation of enterprise systems, the standardisation of information systems, and the availability of outsourcing, all of which eroded the rarity condition that the resource-based view requires. The lesson is instructive: the relationship between a technology and performance is not a structural parameter. It decays as the technology diffuses [21].

*Wave three: digital transformation.* The current wave has been shaped methodologically by the availability of large archival samples and text-based measurement, and empirically by Chinese listed-firm data. A widely used approach measures firm-level digital transformation by counting the frequency of digital-technology keywords in annual report text, and studies employing it have reported that digital transformation improves stock liquidity, raises total factor productivity, and improves input-output efficiency. Against this, the digital servitization literature documents a *digitalisation paradox*, in which manufacturing firms investing substantially in digital offerings fail to capture corresponding revenue growth. The most direct evidence for conditionality comes from work showing that the interaction between digitalisation and high servitization is negative and significant at low-to-moderate levels of digitalisation, turning positive and significant only from moderate to high levels — a functional form that is difficult to reconcile with any linear specification.

The parallel literature on artificial intelligence adoption reproduces the pattern with a distributional twist. Firms investing in AI, measured through employee résumé data, experience higher growth in sales, employment, and market valuation, driven principally by product innovation — but the growth concentrates among larger firms and is associated with rising industry concentration. Adoption itself remains far from universal at the firm level even as individual usage becomes commonplace: United States Census survey data recorded firm AI adoption rising from 3.7 per cent to 5.4 per cent over just three months in late 2023 and early 2024, a rapid increase from a very low base. The gap between near-universal individual usage and low firm-level integration is precisely what a complementarity account predicts: individuals can adopt a chatbot unilaterally; firms cannot reorganise unilaterally.

### 2.4 The Measurement Problem

Text-based measurement of digital transformation has been transformative for the field's empirical scope, and it has a serious weakness that is rarely stated plainly. Counting digital keywords in annual reports measures three things simultaneously: genuine technology deployment, managerial attention to digitalisation, and strategic disclosure aimed at capital markets. It measures a fourth thing barely at all: whether the firm has reorganised around the technology.

This matters for inference in three distinct ways. First, the measure is *increasing in disclosure incentives*, which are themselves correlated with firm size, analyst coverage, and equity issuance plans — all of which independently predict performance. Second, the measure is *conceptually asymmetric*: it captures the numerator of the complementarity gap while remaining silent on the denominator, so that two firms with identical scores may occupy opposite regions of the performance relationship. Third, keyword frequency is *not invariant over time*, because the vocabulary of digitalisation inflates as the concept becomes fashionable, mechanically raising scores independently of any change in behaviour. Studies that have found the estimated effect of digital transformation on outcomes to be systematically smaller in later publication years may be observing this drift rather than a genuine decline in returns.

None of this makes text-based measurement useless. It makes it a measure of one thing that has been used as a measure of another.

### 2.5 The Research Gap

Bringing these threads together, the gap is not an absence of studies. It is an absence of a framework in which the existing studies can all be true. The field has an abundance of estimates of the average effect of a treatment whose effect is theoretically heterogeneous by construction, measured by an instrument that captures one side of a two-sided relationship, in samples drawn from a period in which the measurement instrument itself was drifting. What is required is a specification of the conditioning variable, an operationalisation that existing data can support, and a design that permits the conditional effect to be estimated rather than averaged away. Sections 3 and 4 supply these in turn.

## 3. Theoretical Framework and Propositions

### 3.1 Digital Technology as a General Purpose Technology

The framework begins from the premise that digital technologies — and, in the present period, artificial intelligence in particular — possess the defining characteristics of general purpose technologies: pervasiveness across sectors, continuous improvement over time, and the capacity to spawn complementary innovation in using industries. The economically consequential feature is the third. General purpose technologies deliver their returns not through direct substitution for existing inputs but by enabling co-invention in adopting organisations, and co-invention is costly, slow, firm-specific, and largely intangible.

Consider what a manufacturing firm must actually do to extract value from a manufacturing execution system. It must instrument its equipment; standardise its process definitions so that data from different lines are commensurable; establish data ownership and quality accountability; retrain supervisors to make decisions from dashboards rather than from experience; alter the incentive system so that supervisors are not punished for the variance that measurement newly reveals; and, frequently, redraw the boundary between plant-level and headquarters-level decision rights. Only the first of these is a technology purchase. The remainder are organisational investments, they are expensed, and they take years.

### 3.2 The Complementarity Gap

Let $D_{it}$ denote firm $i$'s stock of digital technology at time $t$, and $C_{it}$ its stock of complementary organisational capital. Following the complementarity logic, firm output is a function in which the cross-partial derivative with respect to $D$ and $C$ is positive: $\partial^2 Y / \partial D\, \partial C > 0$. Define the complementarity gap as

$G_{it} = d_{it} - c_{it}$

Where $d_{it}$ and $c_{it}$ denote $D_{it}$ and $C_{it}$ standardised within industry and year. $G_{it}$ is positive for firms that have acquired digital technology in advance of the organisational capacity to exploit it, negative for firms whose organisational capacity outruns their technology stock, and near zero for firms in balance.

Three properties follow. First, because complementary investment is expensed while its returns are deferred and partly unmeasured, firms with large positive $G$ should exhibit *depressed* contemporaneous measured performance even where long-run returns are positive. Second, because the cross-partial is positive, the marginal return to digital investment is increasing in $C$: identical increments of $D$ produce different performance outcomes at different levels of $C$. Third, because $C$ accumulates slowly and is path-dependent, $G$ is not a choice variable in the short run, which is what makes it a genuine constraint rather than a management failing.

This construction reframes the empirical literature. A study that regresses performance on $D$ alone estimates a coefficient that is a sample-weighted average across the distribution of $C$ in that sample. Two studies drawing samples with different $C$ distributions — different industries, different countries, different periods — will recover different coefficients from the same underlying structural relationship, and will appear to contradict one another while doing nothing of the kind. This is exactly the pattern the meta-analytic evidence reports, with measurement method, industry type, and national economy identified as significant moderators of effect size.

### 3.3 Propositions

3.3.1Proposition 1 (Modest average effect).

*Across a broad sample of firms, the average contemporaneous effect of digital transformation intensity on accounting performance is positive but small in magnitude relative to its dispersion.*

The reasoning is that the sample average is a mixture over the distribution of $C$, combining positive effects among high-$C$ firms with null or negative effects among low-$C$ firms. The prediction is deliberately unambitious: the framework's claim is not that digitalisation fails, but that its average is uninformative. A well-identified estimate that is positive, small, and accompanied by substantial residual heterogeneity is more consistent with the framework than a large positive estimate would be.

3.3.2 Proposition 2 (J-curve dynamics).

*The effect of digital transformation on measured performance is negative or null in the initial post-adoption period and becomes positive thereafter, with the trough occurring earlier for firms with higher pre-existing complementary organisational capital.*

This is the firm-level analogue of the Productivity J-curve. The second clause is the framework's own addition and is the more discriminating test: a pure adjustment-cost story predicts a J-curve for everyone, whereas the complementarity account predicts that the depth and duration of the trough are functions of $C$.

3.3.3 Proposition 3 (Complementarity moderation).

*The effect of digital transformation on firm performance is increasing in complementary organisational capital; equivalently, the effect is decreasing in the complementarity gap.*

This is the framework's central empirical claim and follows directly from the positive cross-partial. It implies a negative coefficient on the interaction between digital intensity and the complementarity gap, and a positive coefficient on the interaction between digital intensity and organisational capital.

3.3.4 Proposition 4 (Environmental contingency).

*The performance effect of digital transformation is larger in environments characterised by higher demand volatility and competitive intensity.*

Digital capability delivers value substantially through faster sensing and reconfiguration. Where the environment is stable, the option value of agility is low and digital investment competes against alternative uses of capital on unfavourable terms. Where the environment is turbulent, the same capability is exercised frequently and its value is realised. This proposition is the point at which the dynamic capabilities lens does independent work in the framework.

3.3.5 Proposition 5 (Distributional divergence).

*Digital transformation widens the performance dispersion within industries, because firms at the productivity frontier possess higher complementary organisational capital and therefore convert digital investment into performance at a higher rate than laggard firms.*

This proposition connects the firm-level framework to an aggregate regularity. Labour productivity at the global frontier has diverged from that of other firms since the 2000s, with the divergence strongest in information and communication technology services — a pattern indicative of winner-takes-all dynamics and of weakening diffusion from the frontier to the rest. Firm-level evidence on artificial intelligence points the same way, with AI-driven growth concentrating among larger firms and associated with rising industry concentration. The complementarity framework supplies a mechanism: diffusion of technology is fast because technology is purchasable, while diffusion of complementary organisational capital is slow because it is not.

Proposition 5 has an uncomfortable corollary. If digital technology is a complement to organisational capital rather than a substitute for it, then subsidising technology adoption without addressing complementary capability should increase, not decrease, the dispersion of firm performance — a point returned to in Section 5.

### 3.4 Summary of the Framework

The framework can be stated compactly. Digital transformation is an investment in one input of a complementary pair. Its measured return depends on the level of the other input, which is intangible, slowly accumulated, and largely unobserved by the empirical literature. The observable consequences are a small and unstable average effect, a J-curve in time whose shape depends on organisational capital, a positive interaction with organisational capital and a negative one with the complementarity gap, amplification under environmental turbulence, and widening dispersion across firms.

**Table 1. Propositions, Empirical Signatures, and Corresponding Tests**

| Proposition | Prediction | Empirical signature | Test |
|---|---|---|---|
| P1 Average effect | Positive but small | Modest main coefficient with large residual dispersion | Two-way fixed effects baseline (§4.7) |
| P2 J-curve | Negative then positive | Trough in distributed-lag path; deeper for low-capital firms | Distributed lags and split-sample event study (§4.9) |
| P3 Complementarity | Effect rises with organisational capital | Positive DT × C interaction; negative DT × G interaction | Interaction and threshold models (§4.7, §4.9) |
| P4 Turbulence | Effect rises with volatility | Positive DT × environmental dynamism interaction | Moderated regression (§4.7) |
| P5 Divergence | Dispersion widens | Coefficient increasing across performance quantiles | Unconditional quantile regression (§4.9) |

Note: DT denotes digital transformation intensity, C complementary organisational capital, and G the complementarity gap.

Table 1 summarises the propositions, the empirical signature each implies, and the specification used to test it. Section 4 develops these in detail.

## 4. Research Design

### 4.1 Empirical Setting

Chinese A-share listed firms over 2013–2024 offer an unusually favourable setting for testing this framework, for four reasons.

First, the period brackets an exceptionally rapid national digitalisation, providing variation in $D$ that is large relative to measurement error. Second, Chinese listed-firm disclosure requirements yield unusually rich data on workforce composition, including the educational attainment and functional classification of employees, which permits construction of a complementary organisational capital measure from archival sources. Third, the Broadband China pilot programme, announced in August 2013 and rolled out to pilot cities in batches between 2014 and 2016, supplies a staggered, geographically assigned infrastructure shock that has been used extensively as a quasi-natural experiment and as an instrument for firm digitalisation. Fourth, and less often noted, the setting includes substantial state ownership, which generates variation in governance and decision-rights structures that is plausibly exogenous to firm-level digital strategy and therefore useful for probing the organisational capital channel.

The setting also has limitations, treated in Section 4.9.

### 4.2 Sample and Data

The initial sample comprises all firms listed on the Shanghai, Shenzhen, and Beijing exchanges over 2013–2024. Following standard practice, I exclude financial and insurance firms, whose accounting structure is not comparable; firms under special treatment or delisting risk warning; firms with negative book equity; and firm-years with missing values on core variables. Continuous variables are winsorised at the 1st and 99th percentiles to limit the influence of outliers.

Financial and governance data are drawn from the CSMAR and WIND databases. Annual report full texts are obtained from the Shanghai, Shenzhen, and Beijing exchange disclosure portals. City-level characteristics, including

Broadband China pilot status and per capita GDP, are taken from the China City Statistical Yearbook. The expected panel is unbalanced, with approximately 4,000–5,500 unique firms and on the order of 40,000 firm-year observations.

### 4.3 Measuring Digital Transformation

The primary measure follows the established text-analytic approach: the natural logarithm of one plus the frequency of digital-technology keywords in the management discussion and analysis section of the annual report, aggregated across four technology families — artificial intelligence, blockchain, cloud computing, and big data — plus a fifth category capturing digital applications.

Given the concerns raised in Section 2.4, three refinements are essential rather than optional.

Year-standardisation. Raw keyword counts are standardised within year to remove the mechanical inflation of digital vocabulary over time. Failure to do this conflates increasing digitalisation with increasing verbosity.

Disclosure-incentive adjustment. The residual from a first-stage regression of raw keyword frequency on total report length, firm size, analyst following, institutional ownership, and an indicator for seasoned equity offerings in the following year serves as a disclosure-adjusted measure. The intuition is to strip out the component of digital language attributable to capital-market communication.

Behavioural cross-validation. Text-based measures are validated against three behavioural proxies: the share of intangible assets whose itemised descriptions reference software, information systems, or data platforms; capitalised software and information-system expenditure scaled by total assets; and the count of granted invention patents in International Patent Classification classes associated with digital technologies. Agreement across measures is reported; where they diverge, results are presented for each.

### 4.4 Measuring Complementary Organisational Capital

This is the design's principal methodological contribution and the point at which it departs from existing practice. Complementary organisational capital $C_{it}$ is constructed as the first principal component of four disclosed indicators:

1. Human capital depth — the share of employees holding a bachelor's degree or above, and separately the share classified as technical personnel.
2. Decision-rights decentralisation — a proxy based on the ratio of parent-company to consolidated employees and assets, capturing the extent to which operational authority resides in subsidiaries rather than headquarters, supplemented by the number of consolidated subsidiaries.
3. Managerial and process investment — administrative expenses net of executive compensation, scaled by operating revenue, capturing expenditure on coordination and process infrastructure.
4. Incentive alignment — the presence and breadth of equity incentive plans extending below the executive tier, capturing the extent to which the firm has altered incentives in a manner consistent with delegated decision-making.

Each indicator is standardised within industry and year before extraction of the principal component. The complementarity gap is then $G_{it} = \tilde{D}_{it} - \tilde{C}_{it}$, with both terms standardised within industry-year so that the difference is interpretable.

The construction is not beyond criticism. Administrative expenses capture waste as well as coordination investment; subsidiary counts reflect tax and regulatory considerations as well as delegation. Robustness checks accordingly report results using each indicator singly, using an equal-weighted index, and using a survey-calibrated subsample where such data are available.

### 4.5 Dependent Variables

Three performance measures are used, corresponding to three theoretical constructs.

Accounting profitability is measured by return on assets, defined as net profit over average total assets, with return on equity as an alternative. This is the measure most directly comparable to the existing literature and the one most exposed to the expensing problem that generates the J-curve.

Productive efficiency is measured by total factor productivity, estimated using the Levinsohn–Petrin approach with intermediate inputs as a proxy for unobserved productivity shocks, and cross-checked against Olley–Pakes estimates. This is the measure most directly relevant to the

Academic Education Publishing House
-AEPH-

complementarity mechanism.

Market valuation is measured by Tobin's Q. Because equity markets in principle capitalise expected future cash flows including returns to intangible investment, valuation-based measures should exhibit a shallower J-curve than accounting measures. The divergence between accounting and market measures is itself a test of the framework rather than a nuisance.

Table 2 summarises the definition and source of each variable.

### 4.6 Control Variables

Controls follow standard practice in the firm-performance literature: firm size (natural logarithm of total assets), leverage, firm age, growth in operating revenue, tangible asset intensity, ownership concentration (shareholding of the largest shareholder), board independence, state ownership, and research and development intensity. City-level per capita GDP and the tertiary-sector share control for regional economic development. All specifications include firm and year fixed effects; specifications reported in the appendix add industry-by-year fixed effects to absorb sector-specific shocks such as differential exposure to the 2020–2022 disruption.

Environmental dynamism, required to test Proposition 4, is measured as the standard error of the regression of industry sales on a time trend over the preceding five years, scaled by mean industry sales, with the Herfindahl–Hirschman index of industry concentration used as a second dimension of environmental structure.

**Table 2. Variable Definitions and Data Sources**

| Variable | Symbol | Definition | Source |
|---|---|---|---|
| Return on assets | ROA | Net profit / average total assets | CSMAR |
| Total factor productivity | TFP | Levinsohn–Petrin estimate; Olley–Pakes as alternative | Author's calculation |
| Tobin's Q | TQ | Market value of equity plus book value of debt, over total assets | CSMAR / WIND |
| Digital transformation | DT | Year-standardised, disclosure-adjusted log keyword frequency | Annual report text |
| Organisational capital | C | First principal component of four disclosed indicators (§4.4) | CSMAR |
| Complementarity gap | G | Standardised DT minus standardised C, within industry-year | Author's calculation |
| Broadband pilot | BB | Indicator equal to one from the year of pilot designation | official policy documents |
| Environmental dynamism | DYN | Scaled standard error of industry sales trend regression | Author's calculation |
| Firm size | SIZE | Natural logarithm of total assets | CSMAR |
| Leverage | LEV | Total liabilities / total assets | CSMAR |
| R&D intensity | RD | R&D expenditure / operating revenue | CSMAR |
| State ownership | SOE | Indicator for ultimate controller being a state entity | CSMAR |

### 4.7 Baseline Specification

The baseline two-way fixed effects specification is

$Y_{it} = \alpha + \beta_1 DT_{it} + \beta_2 C_{it} + \beta_3(DT_{it} \times C_{it}) + \gamma' X_{it} + \mu_i + \lambda_t + \varepsilon_{it}$

where $Y_{it}$ is the performance measure, $DT_{it}$ the disclosure-adjusted digital transformation index, $C_{it}$ complementary organisational capital, $X_{it}$ the control vector, and $\mu_i$ and $\lambda_t$ firm and year fixed effects. Standard errors are clustered at the firm level; two-way clustering by firm and city is reported as a robustness check. Proposition 3 predicts $\beta_3 > 0$. A parallel specification replacing the interaction with the complementarity gap predicts a negative coefficient on $DT_{it} \times G_{it}$.

The baseline is presented as a descriptive benchmark, not as a causal estimate. Digital transformation is chosen, not assigned, and the same unobserved managerial quality that drives digitalisation plausibly drives performance directly.

### 4.8 Identification

Three strategies address endogeneity, each with distinct assumptions.

Strategy one: staggered difference-in-differences around Broadband China. The pilot programme assigned cities to treatment in batches from 2014 to 2016. Firms headquartered in pilot cities experience a reduction in the cost of digital infrastructure that is plausibly exogenous to their own digital strategy. The estimating equation is a staggered difference-in-differences specification with firm and year fixed effects and city-level controls.

Conventional two-way fixed effects estimation of staggered designs is biased when treatment effects are heterogeneous across cohorts or over time, because the estimator uses already-treated units as controls for later-treated units and can assign negative weights to some treatment effects. Given that the framework's central prediction is treatment effect heterogeneity, using an estimator that assumes homogeneity would be self-defeating. Estimation therefore uses the group-time average treatment effect estimator of Callaway and Sant'Anna, which restricts comparisons to not-yet-treated units and permits conditional parallel trends, with the interaction-weighted estimator of Sun and Abraham reported alongside. A Goodman-Bacon decomposition is presented to document the share of identifying variation coming from problematic comparisons under the conventional estimator, and the de Chaisemartin–d'Haultfœuille diagnostic for negative weights is reported.

Pre-treatment parallel trends are assessed through an event-study specification with leads and lags relative to pilot designation, with the period immediately preceding designation omitted as the reference.

Strategy two: shift-share instrumentation. As an alternative source of exogenous variation, I construct a Bartik-style instrument interacting each city's historical fixed-line telecommunications density — measured well before the sample period, and therefore predetermined with respect to contemporaneous firm strategy — with national annual growth in digital technology deployment. The exclusion restriction requires that historical telecommunications infrastructure affects current firm performance only through its effect on current digitalisation. This is a strong assumption. It is weakened, though not eliminated, by controlling for historical city characteristics interacted with year fixed effects, and by conducting the analysis within narrow city-size and region strata.

Strategy three: propensity-score-matched and entropy-balanced samples. Firms in pilot and non-pilot cities are matched on pre-treatment size, leverage, profitability, growth, ownership structure, and industry, with the difference-in-differences estimated on the matched sample. Entropy balancing is used as an alternative that avoids discarding observations while achieving exact moment balance.

Convergence of estimates across strategies with different failure modes is treated as the relevant evidentiary standard. Divergence is reported and discussed rather than resolved by selecting the most favourable specification.

### 4.9 Testing Non-Linearity and Dynamics

4.9.1 J-curve estimation (Proposition 2).

A distributed-lag specification includes digital transformation intensity at $t$ through $t-5$, with the cumulative effect traced across horizons. Separately, firms are split at the median of pre-period complementary organisational capital, and the event-study path is estimated within each group. Proposition 2 predicts a deeper and longer trough in the low-$C$ group.

4.9.2 Threshold estimation (Propositions 3 and 5).

The framework implies a threshold in complementary organisational capital above which digital transformation becomes reliably profitable. Rather than imposing a functional form, I estimate the threshold endogenously using Hansen's panel threshold approach, with complementary organisational capital as the threshold variable and digital transformation intensity as the regime-dependent regressor. The number of thresholds is determined by sequential bootstrap testing.

4.9.3 Dispersion effects (Proposition 5).

Unconditional quantile regression estimates the effect of digital transformation at the 10th, 25th, 50th, 75th, and 90th percentiles of the within-industry performance distribution. Proposition 5 predicts a monotonically increasing coefficient across quantiles. As a complementary test, industry-year dispersion in productivity is regressed on lagged industry-year mean digital intensity.

### 4.10 Robustness and Falsification

Beyond the identification checks already described, the design incorporates the following.

*Alternative measurement.* All principal results are re-estimated using each of the three behavioural digitalisation proxies described in

Academic Education Publishing House

Section 4.3, and using the unadjusted keyword measure, so that readers can assess how much of the result is attributable to the disclosure adjustment.

*Sample perturbation.* Results are re-estimated excluding information and communication technology firms, for which digitalisation is an output rather than an input; excluding state-owned enterprises; excluding the 2020–2022 period; and separately for manufacturing and service sectors.

*Placebo tests.* Pilot designation dates are randomly reassigned across cities 1,000 times and the distribution of placebo estimates compared with the actual estimate. A second placebo assigns treatment to a randomly selected pre-treatment year.

*Falsification.* The framework predicts that the interaction of digital transformation with complementary organisational capital should be substantially stronger for productivity than for outcomes that are largely insulated from process reorganisation. Testing the interaction on such an outcome — for instance, the effective tax rate — provides a falsification test: a significant interaction there would suggest that the organisational capital measure is proxying for general firm quality rather than for complementarity specifically.

*Attrition and selection.* Because delisted and acquired firms exit the panel, results are re-estimated on the balanced subsample and with inverse probability of attrition weights.

### 4.11 Limitations of the Design

Four limitations warrant explicit acknowledgement.

First, listed firms are not representative. They are larger, better governed, and more capital-market-exposed than the population of firms, and small and medium-sized enterprises — where the complementarity gap is plausibly widest — are excluded by construction. The framework's predictions regarding Proposition 5 in particular are therefore tested on a truncated sample, which should bias against finding the predicted dispersion effect.

Second, complementary organisational capital is measured indirectly. The four indicators capture observable correlates of the construct rather than the construct itself, and the principal-component aggregation imposes a particular weighting. Survey-based measurement of management practices would be superior, at the cost of sample size and of the ability to exploit the Broadband China shock.

Third, the Broadband China instrument identifies a specific and somewhat narrow channel — the cost of digital infrastructure — and the local average treatment effect it recovers pertains to firms whose digitalisation responds to infrastructure cost. Firms digitalising for competitive or strategic reasons unrelated to infrastructure cost are not represented in that estimate.

Fourth, the framework's temporal predictions require a long post-treatment window, and the most recent technological wave — generative artificial intelligence deployed at scale from 2023 onward — falls almost entirely outside the window over which J-curve dynamics could be observed. The design can characterise the earlier digitalisation wave thoroughly and the current one only provisionally.

## 5. Anticipated Findings and Implications

### 5.1 Theoretical Implications

If the propositions are supported, the principal theoretical contribution is a reconciliation. The literature's contradictory findings become observations drawn from different regions of a single conditional relationship rather than incompatible claims. This has a methodological corollary that extends beyond this literature: where a treatment's effect is theoretically conditional on an unmeasured moderator, meta-analysis of average effects will document heterogeneity without explaining it, and the appropriate response is to measure the moderator rather than to accumulate further estimates of the average.

A second contribution concerns the resource-based view's application to technology. The reassessment finding that the IT capability–performance link had disappeared by the 2000s was interpreted at the time as evidence that commoditisation had eroded the rarity of IT resources. The framework here suggests a complementary reading: what remained rare was never the technology but the organisational capital, and studies measuring the former were always estimating a relationship whose true argument was the latter. On this reading, the finding of a vanishing effect is not evidence that technology stopped mattering, but evidence that the measurement had drifted further from the construct.

A third contribution is to specify a boundary condition on the dynamic capabilities perspective. That literature has established that reconfiguration capacity conditions the outcome of digital transformation. The framework here makes the conditioning quantitative and directional: the relevant quantity is the *gap* between technology and organisation, which implies that firms can err in both directions, and that a firm whose organisational capital exceeds its technology stock is under-invested in technology in a way that the capability literature, focused on the sufficiency of capability, does not naturally express.

## 5.2 Managerial Implications

The framework's managerial implication is a sequencing rule that runs against prevailing practice. If the cross-partial is positive and organisational capital accumulates more slowly than technology can be purchased, then for a firm below the estimated capability threshold, the marginal return to an additional unit of digital investment is lower than the marginal return to an additional unit of organisational investment — training, process standardisation, delegation of decision rights, data governance. The recommendation is not to digitalise less, but to sequence differently: build the denominator before expanding the numerator.

A second implication concerns evaluation. Because the J-curve implies that measured performance declines before it rises, firms that evaluate digital initiatives on twelve- or eighteen-month accounting outcomes will systematically terminate projects during the trough, and will do so most aggressively precisely where the trough is deepest — that is, in the low-capability firms that had the most to gain from persisting. This generates a self-confirming pattern in which digitalisation appears to work only for firms that were already capable. Evaluation horizons should be matched to the expected shape of the return path, and interim evaluation should rely on operational rather than financial indicators.

A third implication concerns acquisition strategy. If complementary organisational capital is the binding constraint and is slow to build internally, acquiring organisations that already possess it may dominate acquiring technology. The framework predicts that acquisitions motivated by capability transfer should outperform those motivated by technology transfer, which is itself a testable extension.

## 5.3 Policy Implications

The policy implications are less comfortable than the standard narrative permits.

Infrastructure programmes such as Broadband China lower the cost of the technology input. The framework implies that lowering the cost of one input in a complementary pair raises the return to that input most for firms that already possess the other — that is, most for firms at the frontier. Such programmes may therefore be efficiency-enhancing in aggregate while simultaneously widening the productivity distribution. This is consistent with the documented divergence between frontier and laggard firms and with the concentration of AI-driven growth among larger firms. It suggests that infrastructure and adoption subsidies are complements to, not substitutes for, policies that build organisational capability: vocational and managerial training, diffusion of management practices, and support for the process-redesign component of digitalisation rather than only its technology-purchase component.

A second implication concerns measurement in official statistics. If complementary organisational investment is expensed rather than capitalised, then during periods of rapid technological diffusion national accounts will understate both investment and output, with the bias reversing later. Statistical agencies developing measures of intangible investment address precisely this, and the framework suggests that firm-level organisational capital measures of the kind constructed here could inform that work.

A third implication concerns competition policy. If the diffusion of digital technology is fast while the diffusion of complementary organisational capital is slow, then technological catch-up will not by itself restore competitive convergence. Policies aimed at market contestability may be more consequential for the productivity distribution than policies aimed at technology access — a conclusion consistent with evidence that productivity divergence has been most extreme in sectors where pro-competitive product market reform was least extensive.

# 6. Conclusion

This paper began from an empirical embarrassment: after two decades of research, the field cannot say with confidence whether digital transformation improves firm performance. I have argued that this reflects a specification error rather than an evidentiary shortfall. Digital technologies are general purpose technologies whose returns are conditional on complementary organisational assets that are intangible, slowly accumulated, and largely unmeasured in the empirical literature. Estimating the effect of digital transformation without conditioning on complementary organisational capital is estimating one argument of an interaction, and unstable coefficients are the expected result.

The complementarity gap — the standardised distance between a firm's digital technology stock and its complementary organisational capital — supplies a construct that makes this conditionality explicit and, importantly, measurable from disclosure data that already exist. The five propositions derived from it are directional and falsifiable, and several of them would embarrass the framework if they failed: in particular, a J-curve whose depth is unrelated to organisational capital, or an interaction between digital intensity and organisational capital that appears equally in outcomes insulated from process reorganisation, would each constitute meaningful evidence against the account offered here.

The research design set out in Section 4 is demanding. It requires text analysis, principal-component construction of a latent organisational measure, heterogeneity-robust staggered difference-in-differences estimation, endogenous threshold estimation, and quantile analysis, applied to a panel of some tens of thousands of firm-years. Its limitations are real, particularly the restriction to listed firms and the indirect measurement of the central moderating construct. But the alternative — accumulating further estimates of an average effect that theory predicts should be uninformative — has now been attempted at sufficient scale to establish that it does not converge.

Two extensions seem most promising. The first is to carry the framework to the current technological wave: generative artificial intelligence has diffused to individuals faster than any previous general purpose technology while firm-level integration remains comparatively limited, and the complementarity account offers a specific explanation for that divergence which the next few years of data will be able to test. The second is to examine the political economy of the complementarity gap directly — whether firms systematically over-invest in the visible, capitalisable, capital-market-legible technology input relative to the invisible, expensed, organisationally painful complementary input, and if so, whether the incentives of managers and of equity markets are jointly responsible. If the answer is yes, the gap is not merely a constraint that firms face. It is one they choose.

**References**


[1] Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, 17(1), 99–120.

[2] Acemoglu, D., Autor, D., Hazell, J., & Restrepo, P. (2022). Artificial intelligence and jobs: Evidence from online vacancies. *Journal of Labor Economics*, 40(S1), S293–S340.

[3] Bresnahan, T. F., Brynjolfsson, E., & Hitt, L. M. (2002). Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *The Quarterly Journal of Economics*, 117(1), 339–376.

[4] Eisenhardt, K. M., & Martin, J. A. (2000). Dynamic capabilities: What are they? *Strategic Management Journal*, 21(10–11), 1105–1121.

[5] Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2), 254–277.

[6] Kohtamäki, M., Parida, V., Patel, P. C., & Gebauer, H. (2020). The relationship between digitalization and servitization: The role of servitization in capturing the financial potential of digitalization. *Technological Forecasting and Social Change*, 151, 119804.

[7] Bick, A., Blandin, A., & Deming, D. J. (2026). The rapid adoption of generative AI. *Management Science*. Advance online publication.

[8] Melville, N., Kraemer, K., & Gurbaxani, V. (2004). Review: Information technology and organizational performance: An integrative model of IT business value. *MIS Quarterly*, 28(2), 283–322.

[9] Brynjolfsson, E., Rock, D., & Syverson, C. (2021). The productivity J-curve: How intangibles complement general purpose technologies. *American Economic Journal:*

*Macroeconomics*, 13(1), 333–372.
[10] Nambisan, S., Lyytinen, K., Majchrzak, A., & Song, M. (2017). Digital innovation management: Reinventing innovation management research in a digital world. *MIS Quarterly*, 41(1), 223–238.
[11] Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2), 175–199.
[12] Warner, K. S. R., & Wäger, M. (2019). Building dynamic capabilities for digital transformation: An ongoing process of strategic renewal. *Long Range Planning*, 52(3), 326–349.
[13] Wu, F., Hu, H., Lin, H., & Ren, X. (2021). Enterprise digital transformation and capital market performance: Empirical evidence from stock liquidity. *Journal of Management World*, 37(7), 130–144.
[14] Chae, H.-C., Koh, C. E., & Prybutok, V. R. (2014). Information technology capability and firm performance: Contradictory findings and their possible causes. *MIS Quarterly*, 38(1), 305–326.
[15] Lin, Q. (2025). A meta-analytic investigation of digital transformation: Antecedents, consequences, and contingencies. *Journal of Business Research*, 200.
[16] Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British Journal of Management*, 14(3), 207–222.
[17] Zhang, W., & Liu, G. (2023). Digitalization and firm centralization: A quasi-natural experiment based on the "Broadband China" policy. *Finance Research Letters*, 51, 103462.
[18] Brynjolfsson, E., Hitt, L. M., & Yang, S. (2002). Intangible assets: Computers and organizational capital. *Brookings Papers on Economic Activity*, 2002(1), 137–198.
[19] Hanelt, A., Bohnsack, R., Marz, D., & Antunes Marante, C. (2021). A systematic review of the literature on digital transformation: Insights and implications for strategy and organizational change. *Journal of Management Studies*, 58(5), 1159–1197.
[20] Milgrom, P., & Roberts, J. (1990). The economics of modern manufacturing: Technology, strategy, and organization. *American Economic Review*, 80(3), 511–528.
[21] Solow, R. M. (1987, July 12). We'd better watch out. *New York Times Book Review*, 36.