Digital Transformation and Financial Soundness in Life Insurance Companies: Evidence from Text Mining

Xi Chen[1], Ruiqi Wang[2,*]
[1]*Qianhai Reinsurance Co., Ltd., Shenzhen, Guangdong, China*
[2]*Linfen Municipal Development and Reform Commission, Linfen, Shanxi, China*
**Corresponding Author*

Abstract: This study examines whether digital transformation is associated with the financial soundness of life insurance companies. The sample covers Chinese life insurers from 2014 to 2023. A firm-year digital transformation index is constructed from corporate website disclosures using text clustering, Word2Vec embeddings, cosine similarity, and text mining. Financial soundness is measured by the Z-score and the three-year rolling standard deviation of return on assets. Estimates from firm and year fixed-effects models show that a higher lagged digital transformation index is associated with a higher Z-score and lower earnings volatility. The results remain stable when the dependent variable is replaced by the absolute change in the solvency adequacy ratio, the sample period is shortened, recently established and specialised insurers are excluded, and an instrumental-variable specification is used. The association is stronger among large insurers. Channel tests further indicate that digital transformation is linked to higher asset turnover and a larger share of insurance risk, which is consistent with improvements in operating efficiency and the protection-oriented nature of insurance business. These findings suggest that digital investment can support financial resilience, although its payoff depends on an insurer's resources and implementation capacity.

Keywords: Digital Transformation; Financial Soundness; Text Mining; Life Insurance Companies; Two-Way Fixed Effects

1. Introduction

Digital technologies are changing how insurers design products, assess risks, process claims, and manage customer relationships. Their influence also extends to functions that are less visible to policyholders, including asset-liability management, internal control, and capital allocation. In China, the 2022 regulatory guidance on the digital transformation of banking and insurance institutions placed this shift within a broader programme of financial-sector modernisation. For life insurers, digitalisation has therefore become more than an information-technology project. It increasingly shapes how firms organise work and compete.

Whether that shift improves financial soundness is less obvious. Life insurers operate with long-duration liabilities, intertemporal asset-liability mismatches, and cash flows that may remain uncertain for many years. A profitable year does not, by itself, indicate a sound insurer. What matters is whether the firm can sustain earnings, absorb shocks, and meet policyholder obligations over time. Digital tools may help by improving underwriting, pricing, fraud detection, and internal coordination. Yet transformation also requires substantial spending on infrastructure, data governance, process redesign, and specialised staff. The benefits may arrive slowly, while implementation failures can add cost and operational risk immediately.

This tension is sharper in life insurance than in many ordinary service industries. Policies can remain in force for decades, and errors in pricing or risk classification may not become apparent until long after a contract is written. Legacy systems are often linked to large portfolios of existing policies, which makes replacement costly and technically difficult. At the same time, life insurers must coordinate product promises, capital requirements, investment returns, and customer service across long horizons. A technology project that improves sales but weakens data quality or increases model risk may do little for financial soundness. The relevant question is therefore not whether an insurer has adopted digital tools, but whether those tools improve the quality and stability of its operating and risk-management processes.

Existing research reflects this ambiguity. Studies of non-financial firms generally connect digital transformation with productivity, innovation, financing conditions, firm value, and risk taking. The evidence is not uniform, however, because the effect varies with organisational capability, firm size, and the depth of technology-business integration. Insurance research has paid closer attention to InsurTech, pricing, customer service, and operating efficiency, but has produced less direct evidence on the long-run financial soundness of life insurers. Results obtained from manufacturers or commercial banks cannot simply be carried over. Insurance liabilities, prudential regulation, and the central role of risk pooling make the industry distinct.

Two gaps follow from this literature. First, profitability and financial soundness are often treated as interchangeable outcomes even though an insurer can report strong current earnings while remaining exposed to volatile investment returns, weak capital buffers, or an unstable product mix. Second, most evidence on insurance digitalisation concerns a particular application or a cross-sectional efficiency measure. Less is known about whether variation in the broader transformation of individual insurers is associated with changes in soundness over time. A firm-level panel is useful because it allows the analysis to separate persistent differences between insurers from changes that occur within the same insurer.

Measurement presents a second difficulty. Generic digital dictionaries developed for listed companies do not always capture the language of underwriting, policy administration, claims, solvency, and protection-oriented products. This study therefore constructs an industry-sensitive digital transformation index from public texts released by life insurers. Corporate news and information pages are collected with Python, and related digital expressions are identified with text clustering, Word2Vec embeddings, and cosine similarity. The resulting firm-year measure is used in a two-way fixed-effects framework.

Website texts are useful because they cover listed and unlisted insurers and appear more frequently than formal annual filings. They can reveal operating initiatives that are not separately itemised in financial statements. They are nevertheless disclosures, not verified technology inventories. The index may capture managerial attention and communication as well as implementation, so it is interpreted as a measure of disclosed transformation rather than a monetary stock of digital capital.

The empirical analysis covers Chinese life insurers from 2014 to 2023. Financial soundness is captured in two complementary ways: a Z-score, for which a larger value indicates a greater buffer against financial distress, and the three-year rolling standard deviation of return on assets (ROA), for which a smaller value indicates more stable earnings. The baseline estimates point in the same direction under both measures. Digital transformation is associated with a higher Z-score and lower ROA volatility. A series of alternative specifications, including an instrumental-variable approach, yields similar results. The estimates are larger and more precisely measured for large insurers. Additional tests suggest that operating efficiency and the protection-oriented character of insurance business are two relevant channels.

This study contributes in three respects. First, it adapts the measurement of digital transformation to insurance-specific language rather than applying a general corporate dictionary without modification. Second, it shifts attention from short-run profitability to financial soundness, a particularly relevant outcome for institutions funded by long-term policyholder liabilities. Third, it examines two channels grounded in the operating features of life insurance: efficiency in deploying assets and the extent to which business retains its insurance-protection function. The analysis does not imply that technology spending is automatically beneficial. Instead, it shows that the organisational use of digital tools, and the resources available to support that use, matter for the financial outcome.

2. Literature Review and Hypothesis Development

2.1 Digital Transformation and Financial Soundness

Digital transformation can lower the cost of collecting and processing information, improve resource allocation, and reshape production and management processes [1,2]. Empirical work has linked it to firm performance and productivity [3-5], although the relationship depends on the firm's strategy and organisational setting. Cenamor et al. [6] show that the value of digital platform capabilities in small and medium-sized enterprises depends on their fit with broader strategic capabilities. Qi and Cai [7] argue that shallow integration between digital tools and operating activities can limit improvements in financial performance, while Liang et al. [8] report a potentially nonlinear performance effect. These studies caution against treating digitalisation as a uniform input with an immediate return.

It is useful to distinguish three related ideas. Digitisation converts analogue information into machine-readable form. Digitalisation applies digital tools to an existing activity, such as accepting claims through an online portal. Digital transformation goes further by changing how information is shared, how decisions are made, and how products or services are organised. The distinction matters empirically because the purchase of software may raise expenditure without improving coordination, whereas a less visible redesign of data and workflow may have a larger economic effect. Studies that rely on a single investment item can therefore capture a different dimension from studies based on patents, textual disclosure, or integrated indices.

The expected performance effect is also time-dependent. Installation and migration costs are concentrated at the beginning of a project, while learning, data accumulation, and process integration take time. Employees may initially operate old and new systems in parallel, and managers may need to revise responsibilities before faster information becomes useful. These adjustment costs help explain why digital transformation can have weak or even adverse short-run effects yet produce gains after the organisation has adapted. They also imply that firm resources and implementation capacity should influence the observed relationship.

Research on insurance has developed along a related but more industry-specific path. Miao [9] describes digital transformation as an organisational process that must connect technology with workflow, governance, and customer service. Blockchain, big data, artificial intelligence, and cloud computing can alter product design, risk classification, pricing, and claims management [10,11]. Evidence from property insurers suggests that InsurTech has a nonlinear relationship with operating efficiency and that the result differs across ownership and operating structures [12]. Machine-learning methods can also improve risk classification and pricing accuracy [13], while digital customer management may strengthen customer retention and value creation [14]. More recent firm-level evidence finds a U-shaped relationship between digital transformation and total factor productivity among Chinese insurers, highlighting both adjustment costs and longer-run gains [15].

Insurance applications also differ in their connection to the balance sheet. Online distribution mainly changes acquisition and service channels. Underwriting analytics affect the quality of newly written risks. Claims automation influences loss adjustment and customer experience, while systems for asset-liability management influence investment and liquidity decisions. Treating all applications as equivalent can conceal these differences. A broad transformation measure is appropriate for the present question because financial soundness is the joint outcome of several operating and risk-management functions, not of one digital application alone.

Financial soundness differs from ordinary accounting performance. For an insurer, it reflects the capacity to absorb losses, maintain stable operations, and honour policyholder claims. Prior studies have examined capital structure, solvency, product mix, governance, and regulation as determinants of this capacity. Jiang and Liu [16], for example, show that an imbalanced product structure is related to profitability and solvency in listed life insurers. Digital transformation may influence the same outcomes by changing the quality and speed of information, the efficiency of internal processes, and the firm's ability to identify and price risk. It may also impose substantial fixed costs and expose the organisation to transition risk.

The distinction has consequences for measurement. A rise in current profitability may be generated by aggressive product growth or investment exposure and may coexist with greater volatility. Conversely, expenditure on data infrastructure can depress current earnings while strengthening monitoring and future resilience. For this reason, studies based only on ROA or return on equity do not fully answer whether digital transformation makes an insurer more stable. A composite distance-to-distress measure and a measure of earnings volatility capture complementary aspects of soundness and reduce dependence on a single accounting ratio.

On balance, the information and coordination benefits should dominate once digital tools are embedded in operations rather than used as isolated applications. Better data can reduce underwriting errors, strengthen monitoring, and improve the allocation of capital. Automated processes can lower recurring costs and reduce earnings volatility. These arguments lead to the first hypothesis:

The prediction is conditional rather than mechanical. Digital transformation can create new exposures through cyber incidents, model error, third-party service concentration, or poorly governed customer data. These risks do not overturn the hypothesis, but they make implementation quality important. The empirical coefficient is expected to reflect the net result of efficiency and risk-management benefits after accounting for the financial and organisational costs borne by the insurer.

H1: Digital transformation improves the financial soundness of life insurance companies.

2.2 The Operating-Efficiency Channel

The resource-based view treats valuable and difficult-to-replicate organisational resources as a source of sustained advantage [17]. Dynamic-capability theory further emphasises a firm's ability to reconfigure those resources as technologies and markets change [18]. Digital systems can contribute to both processes, but only when data, personnel, and workflows are coordinated.

This coordination is important in life insurance because the value chain is long. Product development, underwriting, policy issuance, policy maintenance, claims, customer service, and back-office administration often rely on separate information systems. Repeated data entry and manual review add delay and cost. Integrated data platforms, automated underwriting, and intelligent claims tools can shorten processing time and reduce routine work. Analytics can also support product allocation and marketing by identifying customer needs more accurately [10,13]. At the management level, faster information flows may improve the use of assets and stabilise cash flow.

Several concrete frictions can be reduced. A common customer record avoids repeated verification across sales, underwriting, and service functions. Rule-based processing can route standard cases quickly while sending unusual applications to experienced staff. Claims systems can match policy information, medical records, and prior transactions before a payment decision is made. Management dashboards can reveal changes in surrender, claims, or premium patterns earlier than periodic manual reporting. None of these improvements depends solely on replacing labour. Their value comes from assigning routine work to systems while directing human attention to cases that require judgement.

Operating efficiency links these changes to financial soundness. When assets generate revenue more efficiently and administrative resources are used with less friction, the insurer has a larger margin with which to absorb shocks. The mechanism is not merely cost cutting. Digital tools may also reduce operational mistakes and allow managers to respond sooner to adverse claims or liquidity developments. Accordingly:

The channel may be stronger for firms that already possess compatible data and skilled employees. An integrated platform is more valuable when historical records are reliable and business units use common definitions. By contrast, fragmented legacy systems can make data cleaning and migration expensive, and an overly rapid rollout can disrupt normal operations. The efficiency hypothesis therefore concerns the realised organisational use of digital transformation rather than technology expenditure in isolation.

H2: Digital transformation improves financial soundness by increasing the operating efficiency of life insurance companies.

2.3 The Insurance-Protection Channel

Risk pooling and compensation are central to insurance [19,20]. In life insurance, the quality of protection-oriented business depends on risk identification, product design, actuarial pricing, and the capacity to manage policies over a long horizon. Digital information can sharpen each of these activities. Large and continuously updated datasets allow insurers to distinguish risk profiles more carefully and align coverage with customer needs [10,13].

The same information can support policyholder management after a contract is sold. Behavioural data and service records may help firms identify changing needs, improve communication, and reduce avoidable lapses [14]. Automated claims screening and fraud detection can shorten legitimate claims while identifying abnormal cases. These applications reduce adverse selection and moral hazard and may improve the composition of claims costs [12,20,21]. A stronger protection orientation can, in turn, reduce exposure to fragile, investment-like products and support more stable insurance cash flows.

Protection-oriented business can affect soundness through both the liability side and the customer relationship. Products designed around clearly identified mortality, morbidity, health, or longevity risks are more closely connected to the insurer's risk-pooling function. Better classification and pricing can reduce cross-subsidisation between heterogeneous policyholders, while more responsive service can improve policy persistence. A stable portfolio of continuing policies supports more predictable premium inflows and reduces the liquidity pressure that can accompany large surrender waves. Digital tools may also help insurers identify unmet protection needs instead of competing primarily through short-term investment features.

These benefits require safeguards. Richer customer data do not automatically produce fairer pricing, and opaque models may reproduce historical bias or exclude groups for reasons that are difficult to explain. Aggressive personalised marketing can also encourage unsuitable products. Data quality, model validation, human review, and consumer protection are therefore part of the channel rather than external concerns. A digitally supported protection strategy should improve the match between risk and coverage without weakening transparency or access.

This channel differs from operating efficiency. It concerns what kind of risk the insurer writes and how well that risk is matched, rather than only how cheaply the firm processes business. The third hypothesis is therefore:

H3: Digital transformation improves financial soundness by strengthening the protection-oriented nature of life insurance business.

3. Research Design

3.1 Sample and Data Sources

The sample consists of Chinese life insurance companies observed from 2014 to 2023. The starting year balances the availability of public digital disclosures with the coverage of financial data. Insurers established in 2019 or later are excluded because they do not provide a sufficiently long operating and disclosure history. The financial variables are assembled from the iFinD database, annual reports, solvency reports, the National Bureau of Statistics, and company websites. Text used to construct the digital transformation index comes primarily from the "Company News" and "News and Information" sections of insurer websites and is collected with Python.

The raw data are screened in five steps. First, the establishment-year observation is removed because it generally covers less than a full accounting year. Second, a dash reported for items such as fair-value gains or losses and exchange gains or losses is coded as zero when the disclosure clearly uses the dash to denote no amount. Third, continuous variables are winsorised at the 5th and 95th percentiles. Fourth, institutions whose principal business is pension trusteeship, occupational annuity management, or another activity outside conventional life insurance and the corresponding solvency framework are excluded. Fifth, firms without a continuous series of usable website texts are removed. The resulting unbalanced panel contains 508 firm-year observations for the principal variables; the full-control regressions use 435 observations because of missing control data.

The panel is deliberately left unbalanced. Requiring every firm to be observed in every year would remove many younger or smaller insurers and could make the sample unrepresentative of the market. Firm fixed effects can use the available within-insurer changes without imposing that restriction. The lower observation count in the controlled models is reported because comparisons across specifications must allow for both the added covariates and the change in estimation sample.

3.2 Variable Measurement

3.2.1 Financial Soundness

Two measures capture different dimensions of financial soundness. The first is the Z-score, a composite indicator of the distance between a firm's current financial position and distress. A larger Z-score denotes greater soundness. The second is the three-year rolling standard deviation of ROA (ROA_SD). It captures the volatility of profitability, so a smaller value indicates more stable financial performance. Using both measures avoids relying on a single accounting outcome.

As an alternative outcome in the robustness analysis, the study uses the absolute rate of change in the solvency adequacy ratio. It is calculated as the absolute value of the current ratio minus the lagged ratio, divided by the lagged ratio. A smaller value indicates less volatile solvency.

3.2.2 Digital Transformation

The digital transformation measure is constructed from insurer-generated website texts. After collecting and cleaning the documents, text clustering is used to identify groups of digital expressions in an insurance setting. Word2Vec embeddings map the vocabulary into a semantic space, and cosine similarity expands the initial set of digital terms to expressions used in related contexts. Text mining then produces a firm-year index, denoted DT, from the identified vocabulary. A larger value indicates more extensive attention to and application of digital transformation in the insurer's public operating disclosures. The regressions use the one-year lag, L.DT, because implementation and financial effects are unlikely to occur within the same reporting period. Lagging the measure also reduces, but does not eliminate, contemporaneous reverse-causality concerns.

Construction proceeds from cleaning to semantic expansion and aggregation. Website pages are stripped of navigation material and other repeated non-article content before the text is assigned to an insurer and year. Clustering is used to separate groups of related expressions rather than assume that every technology term has the same insurance meaning. The Word2Vec model then represents words by the contexts in which they occur. Cosine similarity identifies terms located near the initial digital vocabulary in that semantic space, allowing the dictionary to capture expressions used by insurers even when they differ from the language common in manufacturing or banking. The final index aggregates the resulting digital content at the firm-year level.

The index is designed to capture relative change across insurers and over time. It is not interpreted as the amount of money spent on information technology, the number of systems installed, or a direct quality score. Textual disclosure can rise because a firm has launched several operational projects, because management gives digitalisation greater strategic attention, or because communication practices have changed. Firm fixed effects absorb stable differences in disclosure style, and year effects absorb common shifts in industry language. The lagged specification further requires the disclosed transformation signal to precede the financial outcome.

3.2.3 Control Variables

The controls are firm size (the natural logarithm of total assets), leverage, the logarithm of firm age, the premium-income growth rate, the net-asset ratio, pre-tax profit relative to minimum capital, and indicators for regulatory action, ownership change, and auditor change. These variables capture observable differences in business scale, balance-sheet structure, growth, capital returns, governance events, and regulatory exposure.

Each control may be related to both digital transformation and financial soundness. Large and mature insurers generally possess more data and internal technical capacity, but their legacy systems may also be harder to replace. Leverage and the net-asset ratio describe the balance-sheet buffer available to absorb shocks. Premium growth distinguishes stable development from rapid expansion, while pre-tax profit relative to minimum capital captures the return generated from the regulatory capital base. Regulatory action signals supervisory concern, and ownership or auditor changes may coincide with governance disruption, strategic redirection, or changes in disclosure. Including these variables does not solve every endogeneity problem, but it reduces confounding from observable firm-year conditions.

3.3 Econometric Specification

The baseline equations are:

$$Z_{it} = \alpha_0 + \alpha_1 D_{i,t-1} + \boldsymbol{\gamma}' X_{it} + \eta_i + \tau_t + \varepsilon_{it} \quad (1)$$

$$V_{it} = \beta_0 + \beta_1 D_{i,t-1} + \boldsymbol{\theta}' X_{it} + \eta_i + \tau_t + \nu_{it}. \quad (2)$$

Here, Z is the Z-score, V is ROA_SD, D is the digital transformation index, X is the control vector, i indexes insurers, and t indexes years. Firm fixed effects, η_i, absorb time-invariant differences across insurers, while year fixed effects, τ_t, capture shocks common to all firms. The reported standard errors are clustered. H1 predicts $\alpha_1 > 0$ and $\beta_1 < 0$.

Because D is an index, α_1 and β_1 describe the outcome change associated with a one-unit increase rather than a one-percent increase. This distinction is important when interpreting the magnitude of the estimates. The models do not assume that every digital project produces the same return; the coefficient is an average relationship over the observed range of the index and the insurers included in the sample.

For the channel analysis, the dependent variable is replaced in turn by operating efficiency and the protection-orientation measure:

$$M_{it} = \delta_0 + \delta_1 D_{i,t-1} + \boldsymbol{\phi}' X_{it} + \eta_i + \tau_t + \xi_{it}. \quad (3)$$

M denotes the proposed channel variable. Operating efficiency is measured by asset turnover, while the protection-oriented nature of business is measured by the share of insurance risk. Equation (3) is interpreted as channel evidence rather than a complete decomposition of a causal mediation effect.

4. Empirical Results

4.1 Descriptive Statistics

Table 1 reports descriptive statistics. The Z-score averages 34.88 and varies considerably across insurers, with a standard deviation of 48.29. ROA_SD has a mean of 0.0110, indicating that earnings stability also differs across firms. The lagged digital transformation index ranges from 0 to 8.792 and has a mean of 2.885. This dispersion is useful for identification because it shows that insurers did not move through digital transformation at the same pace.

Table 1. Descriptive Statistics

Variable	Symbol	Observations	Mean	Standard Deviation	Minimum	Maximum
Financial soundness	Z-score	508	34.88	48.29	-13.63	285.2
Three-year rolling ROA volatility	ROA_SD	508	0.0110	0.0137	0.000221	0.109
Lagged digital transformation	L.DT	508	2.885	1.653	0	8.792
Firm size	Size	508	6.086	1.762	1.105	10.98
Leverage	Leverage	508	0.878	0.0943	0.309	1.385
Log firm age	Duration	508	2.328	0.672	0	3.638
Premium-income growth rate	Premium_Growth_Rate	508	28.88	631.4	-1.000	14,231
Net-asset ratio	Netasset	508	0.131	0.125	-0.385	1.128
Pre-tax profit/minimum capital	Preromic	506	-0.0586	0.615	-9.087	1.226
Regulatory action	Regulatory	507	0.181	0.386	0	1
Ownership change	SC	439	0.702	0.458	0	1
Auditor change	Audit	506	0.156	0.363	0	1

The controls reflect the diversity of the sample. The mean leverage ratio is 0.878, consistent with the liability-intensive structure of insurance. Premium growth is particularly dispersed. Its high maximum is concentrated among firms in early expansion and should therefore be read alongside the winsorisation procedure and the controls for age and size. Negative values of the net-asset ratio and pre-tax profit relative to minimum capital indicate that a small number of firm-years were under material financial pressure.

4.2 Baseline Estimates

Table 2 presents the baseline estimates. Columns (1) and (3) include firm and year fixed effects without the time-varying controls. Columns (2) and (4) add the full control set. The coefficient on lagged digital transformation is positive and significant at the 1% level in both Z-score specifications. With controls, a one-unit increase in L.DT is associated with a 9.505-point increase in the Z-score.

The ROA_SD results tell the same story in the opposite direction. The coefficients on L.DT are negative and significant at the 5% level before and after controls are included. In column (4), a one-unit increase in the digital transformation index is associated with a 0.000725 reduction in the three-year rolling standard deviation of ROA. The two outcomes therefore provide mutually consistent evidence for H1. The larger Z-score coefficient after the controls are introduced should not be interpreted as proof of a synergistic effect; it indicates that observable differences across firms were masking part of the conditional association in the parsimonious model.

The joint pattern is economically preferable to relying on either outcome alone. A larger Z-score could reflect a temporary improvement in profitability, while low earnings volatility could describe a firm with consistently weak returns. Observing a larger distance from distress together with more stable ROA is more consistent with resilience. Even so, L.DT is a composite textual index, so the coefficients cannot be read as the financial return to a particular software system or project.

4.3 Robustness Tests

4.3.1 Alternative Measure of Financial

Soundness

The first robustness test replaces the dependent variable with the absolute rate of change in the solvency adequacy ratio. Solvency ratios are defined and disclosed under a common regulatory framework and are widely used to assess insurers' capital position [22,23]. A smaller absolute change indicates a more stable solvency position. As Table 3 shows, the coefficient on L.DT is -0.0275 and significant at the 5% level. Because the number of observations is the same as in the full-control baseline model, the result is not driven by a different estimation sample.

Table 2. Baseline Estimates of Digital Transformation and Financial Soundness

Variable	Z-score (1)	Z-score (2)	ROA_SD (3)	ROA_SD (4)
Lagged digital transformation	7.647***	9.505***	-0.000981**	-0.000725**
	(1.693)	(1.961)	(0.000426)	(0.000353)
Log firm age		25.17		-0.0130**
		(22.73)		(0.00605)
Net-asset ratio		231.4***		-0.0294***
		(13.20)		(0.00368)
Leverage		139.5***		-0.0113
		(41.68)		(0.0140)
Ownership change		-45.58**		0.0203***
		(19.47)		(0.00656)
Pre-tax profit/minimum capital		10.40		-0.0142**
		(9.067)		(0.00555)
Firm size		-3.057		-0.00906***
		(8.742)		(0.00323)
Premium-income growth rate		-0.000545		-1.10e-07
		(0.000647)		(3.58e-07)
Regulatory action		-2.029		-0.00270
		(4.245)		(0.00309)
Auditor change		1.753		-0.00150
		(5.557)		(0.00106)
Constant	12.90**	-154.7**	0.0139***	0.100***
	(4.878)	(61.95)	(0.00123)	(0.0300)
Controls	No	Yes	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes
Observations	508	435	508	435
R^2	0.354	0.483	0.541	0.655

*Note: Clustered standard errors are reported in parentheses. ***, **, and * indicate p < 0.01, p < 0.05, and p < 0.10, respectively. The same convention applies to Tables 3-8.*

This result addresses two concerns at once. It uses a supervisory measure rather than a profitability-based indicator, and it evaluates instability in capital adequacy rather than the level of the firm's accounting return. The negative coefficient means that higher lagged digital transformation is associated with smaller year-to-year movements in solvency adequacy. It does not imply that every change in a solvency ratio is undesirable; instead, it shows that the baseline finding is not confined to the particular construction of the Z-score or ROA_SD.

4.3.2 Shorter Estimation Window

InsurTech became more visible in China around 2017, when the International Association of Insurance Supervisors clarified the concept and investment in insurance technology accelerated [24]. The models are therefore re-estimated for 2017-2023. Table 4 reports a positive Z-score coefficient of 9.577 and a negative ROA_SD coefficient of -0.000859. Both are statistically significant and align with the baseline estimates. Restricting the period reduces the number of observations from the full sample to 358, but it concentrates the analysis in years when digital initiatives were more commonly understood as part of insurance strategy. The similarity of the coefficients to those in the full period suggests that the main result is not produced solely by early years in which the language of digitalisation was uncommon or inconsistently used. The relatively high R-squared values indicate strong explanatory fit within this window, although model fit alone is not evidence of causal identification.

Table 3. Robustness Test Using Solvency Volatility

Variable	Absolute Change in Solvency Adequacy Ratio
Lagged digital transformation	-0.0275**
	(0.0130)
Constant	1.110**
	(0.429)
Controls	Yes
Year fixed effects	Yes
Firm fixed effects	Yes
Observations	435
R^2	0.552

Table 4. Robustness Test for the 2017-2023 Period

Variable	Z-score	ROA_SD
Lagged digital transformation	9.577***	-0.000859**
	(2.078)	(0.000348)
Controls	Yes	Yes
Year fixed effects	Yes	Yes
Firm fixed effects	Yes	Yes
Observations	358	358
R^2	0.527	0.696

4.3.3 Subsample Tests

Two subsamples address concerns about unusually volatile firms. The first retains insurers that have operated for more than three years. Newly established insurers often prioritise rapid expansion, which can produce large changes in financial statements that are not representative of a mature operating model. The second retains conventional life insurers and excludes specialised health and pension insurers. These specialised companies generally have narrower product sets, smaller asset bases, and different regulatory or operating constraints [22]. Table 5 shows that L.DT remains positively associated with the Z-score and negatively associated with ROA_SD in both subsamples. The coefficients are statistically significant at conventional levels. After health and pension insurers are excluded, the coefficient magnitudes decline slightly relative to the baseline estimates, but their signs and inference remain unchanged.

Table 5. Robustness Tests Using Alternative Subsamples

Variable	Z-score: Age > 3 (1)	Z-score: Life Only (2)	ROA_SD: Age > 3 (3)	ROA_SD: Life Only (4)
Lagged digital transformation	10.12***	9.070***	-0.000907***	-0.000722**
	(2.071)	(1.960)	(0.000309)	(0.000322)
Constant	-259.9***	-169.5***	0.0977**	0.0987***
	(69.06)	(61.85)	(0.0393)	(0.0289)
Controls	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes
Observations	411	412	411	412
R^2	0.520	0.484	0.598	0.603

The first restriction shows that the result is not driven by the accounting volatility of companies in their initial growth phase. The second addresses business-model comparability. Specialised health and pension insurers can differ from conventional life insurers in product duration, asset scale, and the type of technical investment they require. Retaining only conventional life insurers produces estimates that remain close to the baseline. The modest decline in magnitude is consistent with some heterogeneity across business types, but it does not alter the conclusion that the association is present in the core life-insurance sample.

4.4 Heterogeneity by Insurer Size

Digital transformation involves fixed costs, long implementation cycles, and uncertain timing of returns. Large insurers may be better placed to absorb these costs and combine technology with existing data, technical staff, and distribution networks. To examine this possibility, the sample is divided at the median of firm size. Coefficient differences are evaluated with Fisher's permutation test based on 1,000 bootstrap replications.

Table 6 reports a clear size gradient. In the Z-score models, L.DT is positive for both groups, but the coefficient for large insurers is more than twice that for small and medium-sized insurers. The empirical p-value for the cross-group difference is 0.045. In the ROA_SD models, the coefficient is negative and significant for large insurers but is not statistically different from zero for the smaller group. The cross-group p-value is 0.050, which is borderline at the 5% threshold. Taken together, the estimates suggest that large insurers convert digital investment into financial stability more effectively.

Table 6. Heterogeneity by Life Insurer Size

Variable	Z-score: Large (1)	Z-score: Small/Medium (2)	ROA_SD: Large (3)	ROA_SD: Small/Medium (4)
Lagged digital transformation	11.73***	5.016*	-0.000884***	-0.000431
	(2.960)	(2.843)	(0.000299)	(0.000818)
Controls	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes
Observations	225	204	225	204
R^2	0.650	0.412	0.621	0.717
Fisher permutation p-value	0.045		0.050	

Several mechanisms could produce this gradient. Large insurers can spread the fixed cost of platforms and cybersecurity across a larger portfolio, train specialised teams, and draw on longer histories of policy and claims data. They may also be better able to integrate digital tools across sales, underwriting, claims, and investment functions rather than adopt separate applications. Smaller firms may gain from flexibility and fewer legacy systems, but limited budgets and dependence on external vendors can constrain implementation. The results do not show that digitalisation is unhelpful for smaller insurers; the positive Z-score coefficient is significant at the 10% level. They indicate that the earnings-stability payoff is less precisely estimated and that scale shapes the conversion of digital activity into financial outcomes.

4.5 Endogeneity and Instrumental-Variable Estimates

Reverse causality is a central concern: financially sound insurers may simply have more resources available for digital projects. The baseline use of lagged DT reduces simultaneity but cannot remove this concern.

The instrumental variable interacts the leave-one-out industry mean of digital transformation with each insurer's attention to digitalisation. Industry DT is calculated after excluding the focal firm. Attention is measured by the frequency of digital keywords in news headlines divided by the total number of headline words. The identifying assumption is that, conditional on controls and fixed effects, this interaction affects financial soundness through the insurer's digital transformation rather than through an independent channel.

The first-stage coefficient in Table 7 is positive and significant at the 1% level. The Kleibergen-Paap rk LM test rejects underidentification (p = 0.0039), and the Kleibergen-Paap Wald rk F statistic of 97.40 exceeds the reported Stock-Yogo 10% critical value of 16.38. In the second stage, instrumented DT raises the Z-score and lowers ROA_SD. The estimates are larger than the within-firm baseline coefficients, so their magnitude should be interpreted cautiously; nevertheless, their signs support the main conclusion.

The diagnostic statistics support instrument relevance, but relevance is only one part of the argument. The exclusion restriction requires that the interaction between leave-one-out industry digitalisation and firm attention affects soundness through the insurer's own transformation rather than through another firm-specific response to industry conditions. That restriction cannot be tested directly. In addition, an instrumental-variable estimate may capture a local effect for firms whose transformation responds strongly to the instrument, which need not equal the average fixed-effects relationship. For these reasons, the second-stage coefficients are treated as a robustness check on direction and statistical significance rather than as the preferred magnitudes for policy calculation.

Table 7. Instrumental-Variable Estimates

Variable	First Stage: DT (1)	Second Stage: Z-score (2)	Second Stage: ROA_SD (3)
Instrument	0.800***		
	(3.03)		
Digital transformation		52.752***	-0.011***

		(2.85)	(-2.85)
Kleibergen-Paap rk LM statistic	8.32		
p-value	[0.0039]		
Kleibergen-Paap Wald rk F statistic	97.40		
Stock-Yogo 10% critical value	{16.38}		
Controls	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes
Observations	435	435	435

Note: t-statistics are reported in parentheses, the p-value is in square brackets, and the Stock-Yogo critical value is in braces.

5. Channel Analysis

Jiang [25] notes that a sequence of regressions should not automatically be treated as a formal causal mediation test. Following that caution, this section asks a narrower question: is lagged digital transformation associated with the two intermediate outcomes proposed in H2 and H3, conditional on the same controls and fixed effects? The estimates are interpreted as evidence consistent with the channels, not as a numerical decomposition of the total effect.

This distinction is important because both intermediate outcomes may respond to broader organisational change. A new data platform can affect asset use, product composition, customer retention, and risk control at the same time, while unobserved management quality may influence several of these outcomes jointly. The regressions therefore test whether the observed direction is consistent with the mechanisms developed in Section 2. They do not establish a complete causal chain from L.DT to the channel variable and then to financial soundness.

5.1 Operating Efficiency

Asset turnover is used to measure operating efficiency. A higher turnover ratio indicates that the insurer generates more operating activity from a given asset base. Table 8 shows that the coefficient on L.DT is 3.489 and significant at the 1% level. The result is consistent with the view that integrated data and automated processes allow resources to move through the organisation more efficiently.

For a life insurer, asset turnover should be interpreted with care because the balance sheet is shaped by long-term reserves and investment assets. The measure does not correspond to inventory turnover in manufacturing. It instead provides a broad indication of how effectively the firm uses its asset base in relation to operating activity. A rise can be consistent with faster policy processing, more efficient use of distribution and service resources, and lower administrative friction, provided that it is not generated by an unsustainable contraction in assets.

This interpretation also fits earlier evidence. Li et al. [5] use asset turnover as an operating-efficiency measure and connect improved efficiency with stronger financial performance. Kong et al. [26] show that shorter cash cycles are associated with higher profitability in Chinese manufacturing firms. Hendricks and Singhal [27] document the financial consequences of disruptions to operations and supply chains, illustrating how operating frictions can affect longer-run performance and risk. Although life insurance does not hold conventional inventories, delays in underwriting, policy servicing, and claims create analogous operational costs. The positive coefficient in Table 8 therefore provides evidence consistent with H2.

Digital transformation can raise this efficiency through several linked changes. Standardised data reduce repeated collection and reconciliation; automated rules accelerate routine underwriting and claims; and shared systems allow service staff to view a policyholder's information without transferring the case across disconnected departments. Faster processing can reduce unit costs, but the financial benefit also comes from fewer errors and earlier identification of unusual transactions. These improvements help explain why the operating outcome may be related to soundness even when the immediate objective of a digital project is customer service rather than cost reduction.

5.2 Protection-Oriented Business

The share of insurance risk is used to measure the protection-oriented nature of a life insurer's business. A higher share indicates that a larger part of the firm's activity is tied to underwriting insurance risk rather than to investment-like business. Because China's second-generation solvency regime entered trial operation in 2015 and was fully implemented in 2016, this test uses the 2016-2023 period.

The shorter period ensures that the measure is constructed within a more consistent solvency framework. It also reduces the channel sample to 393 observations, compared with 421 for operating efficiency. The two coefficients in Table 8 should therefore not be compared mechanically in size. They refer to different dependent variables, scales, and estimation samples.

The coefficient on L.DT is 1.607 and significant at the 5% level. Zhang et al. [28], using the same broad protection concept, find that stronger protection-oriented business is associated with lower investment, surrender, solvency, and liquidity risks. Zhong et al. [29] report that life insurers with a stronger protection orientation can reduce crash risk in their equity holdings, whereas a stronger financial orientation has the opposite implication. Jiang and Liu [16] also relate the share of protection products to the capital structure, profitability, and solvency of listed insurers. These studies clarify why a shift in business composition may matter for soundness. The estimate in Table 8 is therefore consistent with H3.

One plausible interpretation is that better data allow insurers to recognise protection needs and classify risks that would otherwise be difficult to serve. More accurate pricing can make protection products commercially sustainable, while digital policy administration can reduce the service cost of long-duration contracts. Customer analytics may also help firms identify lapse risk and intervene before avoidable surrender, supporting more stable premium flows. The result does not show that investment-linked or savings-oriented products are inherently unsound. It suggests that digital transformation is associated with a business mix that places greater weight on the underwriting of insurance risk, which prior evidence connects to stability.

Table 8. Digital Transformation and Proposed Channels

Variable	Operating Efficiency	Protection-Oriented Business
Lagged digital transformation	3.489***	1.607**
	(0.750)	(0.788)
Controls	Yes	Yes
Year fixed effects	Yes	Yes
Firm fixed effects	Yes	Yes
Observations	421	393
R^2	0.647	0.627

*Note: Clustered standard errors are reported in parentheses. ***, **, and * indicate p < 0.01, p < 0.05, and p < 0.10, respectively.*

6. Conclusions and Implications

Using an unbalanced panel of Chinese life insurers from 2014 to 2023, this study constructs a digital transformation index from company website texts and examines its relationship with financial soundness. The fixed-effects estimates are consistent across two outcomes: digital transformation raises the Z-score and reduces the three-year rolling volatility of ROA. The result survives an alternative solvency-based outcome, a shorter post-2017 window, two subsample restrictions, and an instrumental-variable specification. Size matters. Large insurers display a stronger association, especially for earnings stability.

The size result should be read as evidence about implementation conditions rather than as an argument for concentrating digital investment only among market leaders. Smaller insurers also show a positive Z-score coefficient, but the earnings-volatility estimate is imprecise. Their challenge is to select projects that match their data, staff, and financial capacity. Large insurers, by contrast, must manage the complexity of integration and legacy systems even when they possess greater resources.

The channel analysis offers a practical explanation for these findings. Digital transformation is associated with faster asset turnover and a larger share of insurance risk. The first result points to fewer operational frictions; the second suggests that digital tools may help insurers identify, price, and service protection needs more effectively. Because the analysis does not estimate a full causal mediation model, these results are best read as evidence consistent with the proposed mechanisms.

For insurers, the findings argue against a one-size-fits-all transformation strategy. Large companies can justify investment in integrated data platforms, automated underwriting and claims, and enterprise-wide risk systems. Smaller insurers face tighter budget and talent constraints. Cloud services, modular software, shared industry infrastructure, and carefully selected high-frequency use cases may offer a more credible route than attempting to reproduce the technology stack of a national insurer. In both groups, project evaluation should focus on measurable changes in processing time, error rates, retention, claims quality, and capital use rather than on the number of digital initiatives announced.

Implementation should proceed from business problems rather than from a list of fashionable technologies. An insurer can begin by identifying points at which incomplete information, duplicated work, or slow decisions affect policyholders and financial outcomes. Projects should then be evaluated against a small set of operational and risk indicators before they are expanded. This approach makes it easier to stop systems that do not improve performance and to distinguish genuine transformation from a rise in digital terminology. It also encourages cooperation between actuarial, operating, technology, compliance, and customer-service teams instead of assigning transformation to the information-technology department alone.

Data governance is a practical prerequisite. Common definitions for customers, policies, claims, and risk events reduce reconciliation problems and make analytical results comparable across departments. Access controls, audit trails, model validation, and clear responsibility for third-party services should be built into project design. For smaller insurers, external cloud or software providers may lower fixed costs, but contracts should address continuity, portability, cybersecurity, and the ability to review automated decisions. Cost savings obtained through vendor concentration should not create a new single point of operational failure.

The policy implication is similarly differentiated. Common data standards, secure interfaces, and lawful data-sharing arrangements with health and pension-service providers would lower the fixed cost of digitalisation. Smaller life, health, and pension insurers may benefit from shared technical services and controlled pilot programmes. At the same time, support for innovation should be paired with rules on data classification, model governance, cybersecurity, and consumer protection. Digital risk pricing and targeted marketing deserve particular scrutiny because poor data or opaque algorithms can reproduce bias at scale.

Supervision can focus on outcomes and controls rather than technology labels. Relevant indicators include service continuity, model overrides, data-quality incidents, claim-processing errors, complaint patterns, outsourcing concentration, and the effect of system changes on capital and liquidity. Controlled pilots can be useful where a new application has uncertain risks, but successful testing should not substitute for continuing monitoring after deployment. Shared standards and infrastructure should also be designed so that smaller insurers can participate without surrendering control over commercially or personally sensitive data.

Three limitations remain. Website texts may capture communication strategy as well as actual technology use. The instrumental variable improves identification but its exclusion restriction cannot be tested directly. Finally, the sample is drawn from one national market with a specific solvency regime. Future research could combine textual measures with technology expenditure or system-use data, examine individual digital applications, and test whether the relationship changes under different regulatory frameworks.

Several extensions would strengthen the evidence. Future studies could compare website disclosure with information-technology expenditure, software assets, digital patents, staff composition, or system-use records to separate attention from implementation. Project-level data could identify whether underwriting, claims, distribution, or asset-liability applications contribute differently to financial outcomes. Longer panels would allow researchers to examine adoption costs and delayed returns directly rather than rely only on a one-year lag. Finally, cross-country comparisons could show how data rules, outsourcing markets, product structures, and solvency regimes condition the relationship between transformation and soundness.

The central conclusion is therefore measured. Digital transformation is not a guarantee of financial resilience, and the evidence does not justify indiscriminate technology spending. It is associated with stronger soundness when it is embedded in operating processes and supported by sufficient organisational capacity. For life insurers, the value of digitalisation lies less in the visibility of new tools than in their ability to improve information, execution, and the quality of protection delivered over the long term.

References

[1] Wang H, Feng J, Zhang H, et al. The effect of digital transformation strategy on performance. International Journal of Conflict Management, 2020, 31(3): 441-462.
[2] Liu S C, Yan J C, Zhang S X, et al. Can digital transformation in corporate management improve input-output efficiency? Management World, 2021, 37(5): 170-190 [in Chinese].
[3] Jiang Q P, Liu Y Y, Duan L T. Digital transformation, diversification, and corporate performance. Journal of Technology Economics, 2023, 42(4): 82-96 [in Chinese].
[4] Gao D, Yan Z, Zhou X, et al. Smarter and prosperous: Digital transformation and enterprise performance. Systems, 2023, 11(7): 1-15.
[5] Li X M, Chen Z S, Chen Y Q. Corporate digital transformation and financial performance: Evidence from digital patent data of listed firms. Journal of Technology Economics, 2024, 43(1): 73-87 [in Chinese].
[6] Cenamor J, Parida V, Wincent J. How entrepreneurial SMEs compete through digital platforms: The roles of digital platform capability, network capability and ambidexterity. Journal of Business Research, 2019, 101: 196-206.
[7] Qi Y D, Cai C W. The multiple effects and mechanisms of digitalisation on the financial performance of manufacturing firms. Study and Exploration, 2020(7): 108-119 [in Chinese].
[8] Liang L N, Zhang G Q, Li H, et al. Economic effects of corporate digital transformation: An analysis based on market and financial performance. Modern Management Science, 2022(5): 146-155 [in Chinese].
[9] Miao L. Paths of digital strategic transformation in insurance companies. Insurance Studies, 2019(4): 57-65 [in Chinese].
[10] Xu X. Blockchain and insurance innovation: Mechanisms, prospects and challenges. Insurance Studies, 2017(5): 43-52 [in Chinese].
[11] Huang L Q, Shen Y, Shi H. The impact of InsurTech on the insurance value chain. Economic and Trade Practice, 2018(14): 75-77, 79 [in Chinese].
[12] Xie T T, Zhao X L. The effect of InsurTech on the operating efficiency of property insurers under digital transformation: Evidence from a DEA-Tobit panel model. Research of Finance Development, 2021(3): 53-60 [in Chinese].
[13] Gao Y Q, Meng S W. A two-parameter Tweedie machine-learning model and its actuarial application. Statistical Research, 2024, 41(4): 126-140 [in Chinese].
[14] Lou D Y. The effect of digital customer management on value management in insurance companies. Insurance Theory and Practice, 2024(1): 9-18 [in Chinese].
[15] Wanyan R, Zhao T, Suo L, et al. Digital transformation and total factor productivity in insurance companies: a catalyst or inhibitor? The Geneva Papers on Risk and Insurance - Issues and Practice, 2025, 50(1): 142-184. DOI: 10.1057/s41288-024-00340-1.
[16] Jiang S Z, Liu Y H. The effect of product-structure imbalance on capital structure, profitability, and solvency in life insurers: Evidence from listed insurance companies. Insurance Studies, 2012(3): 45-53 [in Chinese].
[17] Barney J. Firm resources and sustained competitive advantage. Journal of Management, 1991, 17(1): 99-120.
[18] Teece D J, Pisano G, Shuen A. Dynamic capabilities and strategic management. Strategic Management Journal, 1997, 18(7): 509-533.
[19] Arrow K J. Uncertainty and the welfare economics of medical care. American Economic Review, 1963, 53(5): 941-973.
[20] Cummins J D, Weiss M A. Convergence of insurance and banking and implications for financial stability. Philadelphia: Wharton Financial Institutions Center, 2009.
[21] Cummins J D, Rubio-Misas M, Vencappa D. Competition, efficiency and soundness in European life insurance markets. Journal of

Financial Stability, 2010, 6(4): 209-221.
[22] Zhong S M, Zhao G Q. The effects of operating models on the financial condition of life insurance companies: An asset-liability management perspective. Business and Management Journal, 2018, 40(9): 155-172 [in Chinese].
[23] Zhu J G, Su J L, Huang W. Ownership structure, equity regulation and financial performance: Empirical evidence from equity regulation in China's life insurance industry. Accounting Research, 2020(6): 61-74 [in Chinese].
[24] Wanyan R Y, Suo L Y. InsurTech's impact on China's insurance industry. Insurance Studies, 2019(10): 35-46 [in Chinese]. DOI: 10.13497/j.cnki.is.2019.10.003.
[25] Jiang T. Mediating and moderating effects in empirical causal inference research. China Industrial Economics, 2022(5): 100-120 [in Chinese]. DOI: 10.19581/j.cnki.ciejournal.2022.05.005.
[26] Kong N N, Zhang X M, Lv J. The effect of working-capital management efficiency on corporate profitability: Evidence from Chinese listed manufacturing firms. Nankai Business Review, 2009, 12(6): 121-126 [in Chinese].
[27] Hendricks K B, Singhal V R. An empirical analysis of the effect of supply chain disruptions on long-run stock price performance and equity risk of the firm. Production and Operations Management, 2005, 14(1): 35-52. DOI: 10.1111/j.1937-5956.2005.tb00008.x.
[28] Zhang Z C, Guo Z H, Zhang W X, Xiao Y. The impacts of insurance business's protection attribute on performance of life insurance companies. Insurance Studies, 2021(8): 35-51 [in Chinese]. DOI: 10.13497/j.cnki.is.2021.08.003.
[29] Zhong S M, Zhao G Q, Zhang S H. Impact of the operating attribute diversity of life insurance institutions on the stability of financial market. Journal of Finance and Economics, 2021, 47(3): 125-139 [in Chinese].