

Article

Beyond AI Narratives: AI Washing and Organizational Resilience

Yufei Xia * , Jikang Sun, Jiarun Liu, Kun Fang, Huiyi Shi and Na Li

Business School, Jiangsu Normal University, Xuzhou 221116, China; 3020214101@jsnu.edu.cn (J.S.); 2020250053@jsnu.edu.cn (J.L.); 2020240044@jsnu.edu.cn (K.F.); 2020240057@jsnu.edu.cn (H.S.); 2020230109@jsnu.edu.cn (N.L.)

* Correspondence: 6020180093@jsnu.edu.cn; Tel.: +86-15162172700

Abstract

Artificial intelligence (AI) is widely viewed as a technological foundation for organizational resilience. Yet firms may strategically exaggerate their AI-related narratives without corresponding substantive investment. This study examines whether such AI washing is associated with lower organizational resilience. We conceptualize AI washing as a narrative–investment misalignment within organizational systems, in which symbolic AI claims move ahead of substantive AI investment and capability formation. Based on Chinese A-share listed firms during 2010–2024, we develop a firm-level AI washing index by comparing firms’ within-industry ranking in AI disclosure with their within-industry ranking in actual AI investment. AI disclosure is identified from annual reports using a large language model, while actual AI investment is measured through AI-related software and hardware investments. Using double-debiased machine learning, we estimate a significantly negative association between AI washing and organizational resilience. Economically, a one-standard-deviation increase in AI washing is associated with a decline in organizational resilience equivalent to approximately 3.276% of the average annual change in organizational resilience. This estimated pattern remains stable when we employ alternative variable constructions, replace the machine learning algorithms, adjust the cross-fitting folds, use propensity score matching, and further apply a deep instrumental variable strategy. Mechanism tests based on organizational legitimacy provide evidence consistent with legitimacy-related transmission channels, suggesting that AI washing is associated with lower resilience through weakened pragmatic, moral, and cognitive legitimacy under the maintained mediation assumptions. Further analysis reveals an asymmetric pattern: firms whose AI narratives exceed actual investment experience lower resilience, whereas firms whose actual investment exceeds external narratives exhibit higher resilience. The negative estimated association is particularly evident in high-tech industries, enterprises with established bank-firm ties, and enterprises with higher educational heterogeneity in their top management teams. This study advances research on AI disclosure and organizational resilience by showing that symbolic AI narratives can signal system-level fragility when technological claims are misaligned with substantive capability formation.



Academic Editor: Maja Meško

Received: 19 May 2026

Revised: 9 July 2026

Accepted: 10 July 2026

Published: 17 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: AI washing; organizational resilience; double debiased machine learning; large language model; deep instrumental variables

1. Introduction

Amid persistent turbulence in the external environment, organizational resilience has become a critical foundation for sustaining firms’ long-term competitive advantage [1]. Existing studies generally suggest that artificial intelligence (AI) can improve information

processing efficiency, strengthen risk identification, and enhance resource allocation [2]. This view, however, implicitly assumes that firms' AI-related narratives are supported by substantive AI investment and capability formation. In practice, firms may emphasize AI strategies, applications, or technological prospects in external disclosures without making corresponding investments in AI-related software, hardware, data infrastructure, or operational integration. We define this divergence as AI washing: a misalignment between AI narratives and substantive AI investment and capability formation. Such misalignment may reflect symbolic communication and, in some cases, opportunistic incentives to obtain external attention, legitimacy, or valuation benefits. Unlike genuine AI applications, AI washing conceals capability shortfalls and weakens external stakeholders' ability to assess firms' true technological foundations [3,4]. AI may therefore become not only a source of organizational resilience but also a source of organizational fragility when AI-related claims are supported more by symbolic narratives than by substantive AI investment and capability formation.

It is important to distinguish AI washing from several related but conceptually different constructs. AI disclosure refers to firms' textual communication about AI strategies, applications, technologies, or future plans, consistent with recent studies that use textual information to identify and evaluate corporate disclosure content [5]. AI adoption and AI investment, by contrast, refer to substantive commitment and implementation, including AI-related software, hardware, data infrastructure, technical systems, complementary intangible investment, and operational integration [6]. AI washing arises only when these two dimensions become misaligned, especially when AI-related narratives move ahead of substantive AI investment and capability formation. Thus, AI washing is not equivalent to AI disclosure itself, nor does it imply genuine AI adoption. It is a specific form of technology washing, analogous to greenwashing in its reliance on a gap between external claims and substantive actions, but rooted in the complexity, opacity, and weak external verifiability of AI technologies [7]. This distinction is central to our argument because organizational resilience depends not on the volume of AI-related communication, but on whether such communication is supported by credible capability formation, stakeholder trust, resource mobilization, and coordinated adaptation under uncertainty [8].

China provides an important setting for examining this issue. First, Chinese firms have long operated under economic transformation and industrial upgrading [9]. As technological change accelerates, business uncertainty also rises. Organizational resilience has therefore become essential for firms' continued survival. Second, China has actively promoted the extensive incorporation of AI into real-sector economic activities in recent years [10]. Policy support has strengthened firms' external incentives to adopt AI and has also increased capital market attention to AI narratives. Under this background, firms are more likely to use AI-related disclosure to respond to external expectations. Third, annual reports of Chinese listed firms are relatively systematic. Data on AI textual disclosure and AI-related investment are therefore more accessible. This provides a solid basis for identifying the divergence between AI narratives and actual investment. Accordingly, using Chinese A-share listed firms as the empirical context helps reveal the real consequences of AI washing on organizational resilience.

Existing studies usually treat AI as a technological foundation for organizational resilience [11]. The core logic is that AI can improve information processing efficiency and help firms identify external risks more quickly [12]. Firms can therefore allocate resources more effectively and restore operations more rapidly after shocks. This logic, however, rests on a critical premise. AI must be translated into real capability rather than remaining at the level of external narrative. When firms use AI mainly as a strategic communication tool, AI may shift from a source of resilience to a source of resilience risk. The harm of AI washing

lies not only in overstating technological capability. More importantly, it creates a false sense of technological security [13]. External stakeholders may overestimate firms' digital capability. Firms themselves may also misjudge their technological boundaries. As a result, systematic distortions are more likely to emerge in resource allocation and risk response.

From this perspective, AI washing offers a reverse lens for understanding organizational resilience. A genuine AI application can strengthen firms' capacity to cope with uncertainty. AI washing, by contrast, may weaken firms' actual preparedness for uncertainty. Firms may project a stronger technological image on the surface, while their shock response capacity does not improve accordingly. Once external shocks occur, the gap between narrative and capability becomes more visible. Stakeholders may then reassess the credibility of corporate disclosure. Firms themselves may also face resource misallocation and delayed action. As a result, AI washing not only fails to strengthen organizational resilience but may instead weaken firms' ability to absorb shocks and restore operations [14].

Building on the foregoing logic, AI washing goes beyond being a generic form of disclosure bias. It reflects a systematic mismatch between technological narratives and substantive capability. If such a mismatch weakens organizational resilience, existing research on the positive effects of AI needs to distinguish more clearly between genuine AI capability and symbolic AI narratives. Accordingly, this study examines whether narrative capability mismatch translates into resilience loss and explains the internal transmission channels through which this effect operates from an organizational legitimacy perspective.

Empirical evidence indicates that AI washing is significantly negatively associated with firms' organizational resilience. Mechanism tests indicate that the estimated relationship is consistent with legitimacy-related transmission channels, specifically the pragmatic, moral, and cognitive dimensions of organizational legitimacy. Additional analyses reveal an asymmetric pattern. Further analysis shows an asymmetric pattern. Firms whose AI narratives exceed actual investment experience lower resilience, whereas firms whose actual investment exceeds external narratives show higher resilience. Heterogeneity tests further indicate that the estimated relationship varies with institutional environment, industry characteristics, financing relationships, and management team features. Relative to the existing literature, this study extends the field in the following three ways.

First, our analysis contributes to the literature on AI disclosure and AI adoption by developing an LLM-based, firm-level measure of AI washing that captures the mismatch between what firms say about AI and what they substantively invest in. Existing studies have begun to examine AI-related disclosure and AI adoption, but they often treat AI narratives as indicators of firms' technological orientation or capability [5,15]. By contrast, this study focuses on the gap between AI disclosure and substantive AI investment. We identify AI-related disclosure from corporate annual reports using LLMs and benchmark it against firms' actual AI investment. Specifically, we construct an AI washing index based on the within-industry percentile spread between a firm's ranking in AI disclosure and its ranking in substantive AI investment. This design allows us to distinguish symbolic AI narratives from substantive AI adoption and provides a more direct measure of narrative-investment misalignment in the AI context.

Second, this study extends research on technology washing and greenwashing to the context of AI and emerging technologies. Prior studies on greenwashing emphasize the gap between environmental claims and environmental performance, while recent work on technology washing highlights symbolic claims around complex and weakly verifiable technologies [3,13]. AI washing shares this general logic, but it also has distinctive features. AI technologies are technically complex, externally difficult to verify, and often surrounded by strong market and policy attention. These features make it easier for firms to construct

persuasive AI narratives even when substantive capability formation lags behind. By examining AI washing, this study shows how the logic of washing operates in an emerging technology setting and clarifies why AI-related narrative exaggeration has organizational consequences that differ from traditional greenwashing.

Third, this study contributes to the organizational resilience literature by showing how AI washing can weaken firms' ability to withstand and recover from shocks. Existing studies usually regard AI as a technological foundation for organizational resilience and emphasize that AI can improve information processing, risk identification, and shock response capacity [16,17]. We extend this perspective by shifting attention from AI capability itself to the credibility of AI-related capability claims. When AI narratives exceed substantive AI investment, firms may distort stakeholder expectations, weaken resource mobilization, and undermine governance trust and coordinated adaptation. The mechanism is consistent with the view that symbolic AI narratives may erode the organizational foundations of resilience. In doing so, the study offers a reverse perspective on the AI-resilience relationship: AI may strengthen resilience when supported by substantive investment and capability formation, but symbolic AI narratives may become a source of organizational fragility.

2. Literature Review, Theoretical Analysis, and Hypothesis Development

2.1. Conceptual Foundations of AI Washing and Organizational Resilience

2.1.1. Definition and Economic Repercussions of AI Washing

Following the rise of greenwashing, AI washing has emerged as a salient form of strategic disclosure in the context of frontier technologies. Greenwashing refers to firms' attempts to create an environmentally responsible image through misleading or weakly substantiated claims [18]. By analogy, AI washing describes firms' strategic exaggeration of AI-related capabilities, applications, or technological prospects without corresponding substantive investment or implementation [7,19]. Such behavior is closely tied to the opacity of AI technologies. It is also reinforced by the difficulty external stakeholders face in assessing firms' actual AI capabilities. From the perspective of narrative economics, technology narratives can shape expectations and influence economic behavior [20]. AI washing, therefore, reflects not only disclosure distortion but also firms' active construction of symbolic technological narratives.

The economic consequences of AI washing first appear in capital markets. AI-related narratives can attract investor attention and generate short-term valuation premiums. Prior studies suggest that AI announcements may trigger positive market reactions when investors perceive them as signals of technological transformation [21]. However, when firms with limited technological foundations issue similar AI claims, the information environment becomes more opaque. Credible and superficial AI signals coexist. This raises investors' costs of identifying genuine technological capability and may produce a lemons market effect [22]. In the long run, firms that rely on narrative-based valuation without substantive capability support may face greater valuation uncertainty and higher reputational risk. AI washing can therefore weaken market discipline rather than improve informational efficiency.

AI washing may also affect firms' internal strategic behavior. To maintain an image of technological leadership, firms may allocate managerial attention and organizational resources toward impression management. Such efforts can crowd out substantive capability building. They may also delay investment in data infrastructure, digital talent, and operational integration. More importantly, AI washing can create a false sense of technological readiness. Managers may overestimate the firm's adaptive capacity. External stakeholders may also misjudge the firm's ability to respond to shocks. This mismatch between technological narrative and real capability can generate resource misallocation

and organizational inertia. As a result, AI washing may undermine the very capabilities that support resilience.

Despite growing interest in AI washing, existing evidence remains limited. Prior research has mainly focused on its conceptual foundations, market reactions, and financing consequences [3,13]. Much less is known about whether AI washing affects how firms absorb disruptions, maintain operations, and recover after external shocks. This gap is important because organizational resilience is shaped not merely by resource endowments but also by credible information, stakeholder trust, and coordinated response. If AI washing damages these foundations, it may turn AI from a source of resilience into a source of vulnerability. To fill this gap, this study investigates the mechanisms through which AI washing weakens organizational resilience.

2.1.2. Organizational Resilience Definition and Determinants

Against the backdrop of rising environmental uncertainty, organizational resilience has become an important analytical lens for understanding how firms respond to shocks and achieve adaptive adjustment. Existing studies have examined its determinants from multiple perspectives. Overall, organizational resilience depends on a firm's internal capability base. It is also shaped by organizational processes and the external environment. Accordingly, this section reviews the determinants of organizational resilience from three dimensions, namely internal capabilities, organizational mechanisms, and external conditions.

Internal capability accumulation provides an important foundation for organizational resilience. The entry of national industrial investment funds can ease financing constraints and support technological upgrading, thereby strengthening firms' capacity to withstand external risks [23]. Higher ESG performance can improve governance quality and stabilize stakeholder relationships. This helps firms maintain operational continuity under uncertainty [24]. Taken together, internal factors enhance organizational resilience by strengthening resource support and improving governance quality.

Organizational mechanisms and network structures also shape resilience. Higher supply chain concentration can weaken corporate shock-resistance by restricting financing access, constraining innovation capacity, and dampening firms' risk-taking incentives [25]. This suggests that external network stability is important for organizational resilience.

External conditions further influence the formation of organizational resilience. The rollout of national big-data pilot-zone policies can improve digital infrastructure and enhance firms' ability to adapt to external shocks [26]. More favorable institutional environments can also improve information availability and support firms' operational adjustment. Overall, external conditions enhance organizational resilience by improving the resource and information environment in which firms operate.

2.2. Theoretical Analysis and Hypothesis Development

When firms disclose information on emerging technologies, the effect of such disclosure depends on whether external narratives are supported by substantive investment and capability formation. Under information asymmetry, credible disclosure can reduce uncertainty and strengthen stakeholder recognition when it is backed by sustained investment and observable organizational action. By contrast, when firms rely mainly on symbolic AI narratives without corresponding substantive investment, disclosure may evolve into AI washing. In this study, AI washing refers to a misalignment between AI narratives and substantive AI investment and capability formation. This definition is consistent with recent studies that conceptualize AI washing as a gap between firms' AI-related claims and their underlying technological substance [3,13]. This misalignment matters because

AI technologies are complex, difficult for external stakeholders to verify, and often surrounded by strong market and policy attention. As a result, stakeholders may initially treat AI-related narratives as signals of technological progress, even when actual AI capability formation remains limited. Building on this logic, Figure 1 presents the theoretical framework through which AI washing affects organizational resilience, including the proposed pathways involving pragmatic, moral, and cognitive legitimacy.

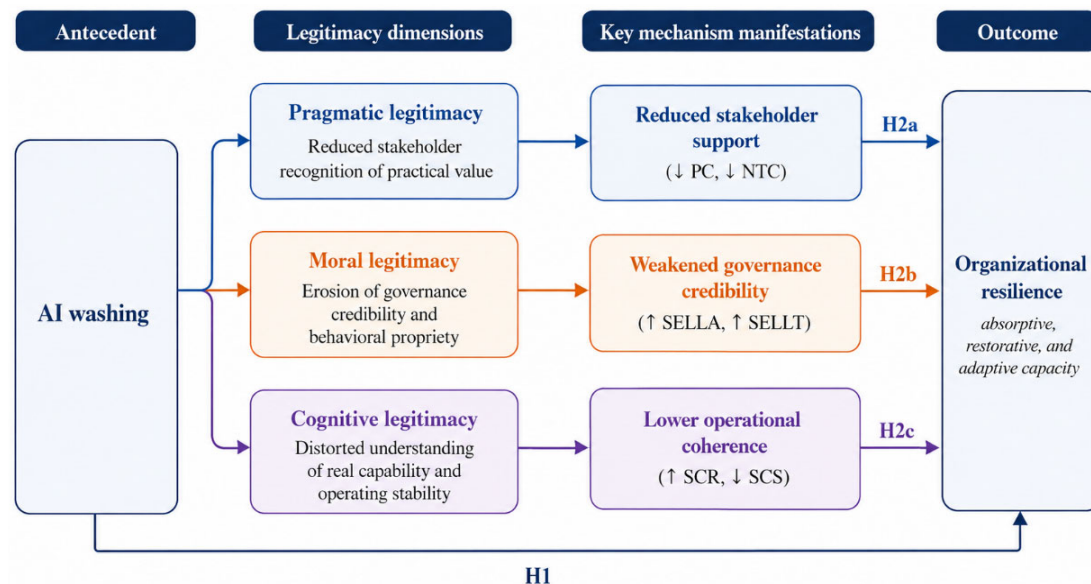


Figure 1. Mechanism through which AI washing affects organizational resilience. The blue, orange, and purple pathways represent pragmatic, moral, and cognitive legitimacy, respectively. The symbols ↓ and ↑ indicate a decrease and an increase, respectively. H1 denotes the direct relationship, whereas H2a–H2c denote the three legitimacy-related pathways.

2.2.1. AI Washing, Legitimacy Loss, and Organizational Resilience

Organizational legitimacy provides a theoretical bridge between AI washing and organizational resilience. Legitimacy refers to stakeholders' generalized perception that a firm's actions are desirable, proper, or appropriate within a socially constructed system of norms, values, and beliefs. Following Suchman [27], we distinguish three dimensions of legitimacy: pragmatic legitimacy, moral legitimacy, and cognitive legitimacy. Pragmatic legitimacy reflects stakeholders' assessment of whether the firm can generate practical value and provide expected benefits. Moral legitimacy reflects whether the firm's conduct is viewed as appropriate, responsible, and consistent with accepted norms. Cognitive legitimacy reflects whether the firm's strategy and behavior are understandable, coherent, and taken for granted by stakeholders.

The link between legitimacy and resilience can be further explained through resource dependence theory and the dynamic capabilities perspective. Resource dependence theory suggests that firms rely on external stakeholders for critical resources, including capital, credit, supply-chain cooperation, and market support [28]. Organizational resilience, therefore, depends not only on internal resources but also on whether stakeholders are willing to continue providing support when the firm faces uncertainty. The dynamic capabilities perspective further emphasizes that resilient firms must sense environmental changes, mobilize resources, and coordinate adaptive responses [29]. These processes require stakeholder trust, governance credibility, and shared understanding of the firm's strategic direction.

AI washing can weaken these foundations. When AI narratives exceed substantive AI investment, stakeholders may reassess the firm's value creation capacity, question the

propriety of managerial behavior, and find the firm's technological strategy less coherent. This process can reduce resource support, increase governance frictions, and impair coordinated adaptation. Therefore, AI washing may affect organizational resilience more directly than outcomes such as stock prices, borrowing costs, or employee turnover, because resilience requires the joint operation of resource mobilization, stakeholder coordination, and adaptive capability formation under disruption. Prior studies generally suggest that AI can strengthen organizational resilience by improving information processing and shock response capacity [16,17]. In contrast, this study argues that AI-related narratives may weaken resilience when they are not supported by substantive investment and capability formation. Accordingly, this study proposes the following hypothesis:

H1. *AI washing negatively affects organizational resilience.*

2.2.2. Pragmatic Legitimacy

From the perspective of pragmatic legitimacy, the negative effect of AI washing first appears in stakeholders' doubts about the firm's value creation capacity. Pragmatic legitimacy is built when stakeholders believe that the firm can create tangible benefits and sustain mutually beneficial relationships [27]. Genuine AI investment may improve operating efficiency, strengthen decision quality, and enhance the firm's long-term growth potential. AI washing, by contrast, signals that the firm may be using AI narratives without sufficient substantive investment. Once this mismatch becomes salient, stakeholders may reduce their assessment of the firm's technological capability and future value creation.

This logic is closely related to the empirical indicators used in our mechanism analysis. Patient capital and net trade credit both reflect stakeholders' willingness to provide stable and relationship-based resources to the firm. Patient capital captures long-term financial support from investors who are willing to tolerate short-term fluctuations in exchange for long-term value creation. Prior research shows that long-horizon investors can reduce short-termism and support firms' long-term investment decisions [30]. Net trade credit reflects the trust and support embedded in supplier–customer relationships because suppliers and business partners are more willing to extend trade credit when they perceive the firm as reliable and value-creating. Trade credit is especially relevant under information asymmetry because suppliers often possess private information about buyers and provide financing when they have confidence in the buyer's prospects [31]. If AI washing weakens pragmatic legitimacy, investors and trading partners may become less willing to provide these forms of resource support. The firm's resource buffer then contracts, making it harder to absorb shocks and restore operations. Therefore, this study proposes the following hypothesis:

H2a. *AI washing weakens organizational resilience by reducing pragmatic legitimacy.*

2.2.3. Moral Legitimacy

From the perspective of moral legitimacy, the negative effect of AI washing is reflected in stakeholders' reassessment of the propriety of corporate conduct. Moral legitimacy depends on whether stakeholders view the firm's behavior as responsible, appropriate, and consistent with accepted governance norms [27]. AI washing is not merely a disclosure gap. When firms emphasize AI strategies and applications without corresponding substantive investment, stakeholders may interpret such behavior as opportunistic use of technological hype. This interpretation can weaken trust in corporate disclosure and raise doubts about managerial integrity.

Opportunistic executive share reductions provide an empirical indicator of weakened moral legitimacy. Executive share reductions are not necessarily problematic in themselves, but opportunistic reductions may signal that managers use information advantages or

market narratives for private benefit. Prior research links insider trading behavior to earnings quality and managerial information advantages, suggesting that insider trades can reveal concerns about managerial incentives and reporting credibility [32]. When AI washing coexists with opportunistic share reductions, stakeholders are more likely to view managerial behavior as self-serving and inconsistent with responsible governance. This damages moral legitimacy and may increase monitoring pressure, internal governance conflict, and coordination costs. Since organizational resilience depends on credible decision-making and timely internal coordination during disruptions, weakened moral legitimacy can reduce the firm's ability to respond effectively to external shocks. Therefore, this study proposes the following hypothesis:

H2b. *AI washing weakens organizational resilience by damaging moral legitimacy.*

2.2.4. Cognitive Legitimacy

From the perspective of cognitive legitimacy, the negative effect of AI washing is reflected in a decline in the coherence and intelligibility of the firm's strategy. Cognitive legitimacy concerns whether stakeholders can understand the firm's behavior as natural, coherent, and consistent with its organizational capabilities [27]. Genuine AI transformation usually involves consistency among technological investment, business processes, data infrastructure, and operational scenarios. AI washing creates a cognitive gap between what the firm says and what it substantively does. This gap makes it harder for stakeholders to understand whether the firm possesses real AI capability or is merely using AI rhetoric for symbolic purposes.

Supply-chain risk and supply-chain resilience are relevant empirical indicators of cognitive legitimacy because they capture how external partners understand and evaluate the firm's reliability in interorganizational operations. In this study, supply-chain resilience is not used as an alternative measure of overall organizational resilience. Rather, it captures the stability and adaptive capacity of the firm's supply-chain relationships. Supply chains require shared expectations, credible information, and coordinated responses among firms [33]. When AI washing increases uncertainty about the firm's true capability and strategic direction, supply-chain partners may find it more difficult to assess the firm's reliability. This can increase supply-chain risk and weaken supply-chain resilience. At the same time, the firm itself may face greater informational noise and strategic ambiguity, which constrains environmental sensing and adaptive adjustment. Because organizational resilience depends partly on the ability to interpret risks and coordinate responses across organizational boundaries [34], weakened cognitive legitimacy can reduce resilience. Therefore, this study proposes the following hypothesis:

H2c. *AI washing weakens organizational resilience by undermining cognitive legitimacy, as reflected in higher supply-chain risk and lower supply-chain resilience.*

3. Empirical Design

3.1. Data Sources

We employed a comprehensive firm-year dataset covering Chinese Main Board companies over the period 2010–2024. Information on AI application terminology was hand-collected from firms' annual reports, while the remaining variables were obtained from the CSMAR and WIND databases.

To improve data reliability, sample representativeness, and the credibility of the empirical evidence, we implemented a systematic data cleaning and preprocessing procedure. We began with 55,542 firm-year observations for Chinese A-share listed firms from 2010

to 2024. The first screening step excluded 15,073 observations involving ST, ST, and PT firms, financial sector firms, and firms with fewer than three full fiscal years since listing, leaving 40,469 observations. We then excluded 4705 observations with missing values in the explanatory variable, dependent variable, and control variables. The final sample contains 35,764 firm-year observations. Furthermore, the explanatory variables and outcome variables were lagged by one period. Finally, to reduce the influence of outliers, continuous variables were winsorized at the 1% level.

3.2. Variable Definition

3.2.1. AI Washing

This study constructs a measure of AI washing that captures the gap between firms' AI-related narratives and substantive AI investment. Conceptually, this approach follows the broader washing literature, which identifies washing through the inconsistency between symbolic external claims and substantive organizational actions or performance [35,36]. In the AI context, recent studies similarly conceptualize AI washing as a divergence between firms' AI-related claims, disclosure strategies, or technological narratives and their underlying technological capabilities, substantive actions, or investment foundations [3,13,37]. We operationalize this misalignment by comparing firms' AI-related disclosure intensity with their substantive AI investment. Accordingly, the *AIW* index is designed to capture the relative degree of narrative-investment misalignment rather than stakeholders' direct ability to verify each AI-related statement.

Specifically, we measure this divergence by calculating the difference in within-industry percentile rankings between the two dimensions. To begin, we develop a dictionary of AI-related keywords. Following Gilardi, et al. [38], we apply LLM-based text recognition to extract AI-related sentences from corporate annual reports, which is used to capture the intensity of firm-level AI disclosure. To validate our AI disclosure extraction, we manually audited 3000 randomly sampled sentences from annual reports. Two research assistants independently coded the sample, and a supervising reviewer resolved disagreements. The inter-coder agreement was satisfactory, with a Cohen's kappa of 0.771. We then compared several LLMs against this benchmark and found that Qwen-max performed best, with a precision rate of 0.9557 and an F1 statistic of 0.9540. These results support the reliability of our AI disclosure identification procedure. We present the technical details and validation procedure of the LLM-based measurement in Appendix A. Appendix A also reports the keyword dictionary, prompt design, classification criteria, representative LLM outputs, manual verification procedures, and estimation scripts to improve transparency and reproducibility. However, they should be interpreted as validation of the disclosure identification step rather than full validation of the broader AI washing construct. We measure firms' actual AI investment by the ratio of the firms' total investment in AI software and hardware projects to total assets. Drawing upon both measures, this study constructs the AI washing index through the following procedure:

First, we compute the industry-specific percentile ranks for each of the two indicators.

$$P_{LLM,it} = \frac{Rank(AI_LLM_{it}) - 1}{N_{industry,t} - 1}, \quad (1)$$

$$P_{Invest,it} = \frac{Rank(AI_Invest_{it}) - 1}{N_{industry,t} - 1}, \quad (2)$$

where $Rank(AI_LLM_{it})$ and $Rank(AI_Invest_{it})$ represent the firm's ranking within its industry in year t based on $P_{LLM,it}$ and $P_{Invest,it}$, respectively (a higher ranking indicates a larger value). $N_{industry,t}$ denotes the overall count of enterprises in that industry in year t . For industries with only one firm, the percentile rank is defined as 0.5. Notably, we also

use AI-related patents and AI-related recruitment as proxies of firms' actual AI investment, which is discussed in Section 4.3.1.

Based on the percentile rankings above, the AI washing index is constructed as follows:

$$AIW_{it} = (P_{LLM,it} - P_{AI,it}) / 100. \quad (3)$$

The index ranges from -0.01 to 0.01 because the difference between the two within-industry percentile ranks is divided by 100. A larger value of AIW_{it} implies that the firm's AI focus in its information disclosure significantly exceeds its actual level of AI investment, implying a more severe AI washing behavior. Conversely, a smaller value of AIW_{it} suggests that the firm's actual AI investment level surpasses the extent of its AI-related disclosure. An AIW_{it} value close to 0 indicates that the firm's AI information disclosure aligns with its actual level of AI investment.

To distinguish substantive AI-related investment from general digitalization investment, we classify investment items according to the technological content and application context of the underlying software and hardware descriptions, following the criterion of the CSMAR database. An investment item is coded as AI-related only when its description explicitly refers to AI technologies, AI-enabled systems, or AI-specific application scenarios, such as artificial intelligence algorithms, machine learning systems, intelligent recognition systems, intelligent decision-making platforms, AI servers, data training platforms, or AI-based production and service applications. By contrast, general investments in information systems, office software, enterprise resource planning platforms, enterprise management systems, automation equipment, or broad digital transformation are not classified as AI investments unless the disclosed description clearly indicates an AI function or AI-specific application. This rule helps ensure that our investment proxy captures substantive AI-oriented investment rather than general IT, automation, or digitalization expenditure. Detailed keyword examples, exclusion criteria, and coding rules are reported in Appendix A.

3.2.2. Definition of Organizational Resilience

Organizational resilience refers to a firm's capacity to absorb disruptions, maintain operational continuity, recover from adversity, and adapt to changing environments [39,40]. Prior studies emphasize that organizational resilience is a multidimensional construct that can be reflected in firms' sustained operating performance and financial stability over time. Following Ortiz-de-Mandojana and Bansal [41] and Sajko, Boone and Buyl [8], this study constructs a composite organizational resilience index from two dimensions: medium-term operating recovery and financial stability.

The medium-term operating performance dimension captures whether a firm can sustain recovery and expansion over a medium-term window. It is measured by the three-year cumulative increase in operating revenue:

$$MTPerf_{i,t} = Revenue_{i,t} - Revenue_{i,t-3}, \quad (4)$$

where $Revenue_{i,t}$ represents firm i 's operating revenue in year t . To mitigate scale-related differences across firms and reporting units, $MTPerf_{i,t}$ is converted into 100 million yuan before normalization. A larger value indicates stronger sustained performance capacity.

The financial stability dimension captures whether a firm can maintain stability under external uncertainty. Following prior resilience research, we measure this dimension using the volatility of monthly stock returns. Lower stock-return volatility indicates stronger financial stability and a greater ability to withstand external uncertainty. Although this component is market-based, it complements the operating-recovery dimension by

capturing the stability side of resilience. Thus, the composite index combines a medium-term operating dimension with a financial-stability dimension and is intended to capture firms' broader absorptive, restorative, and adaptive capacity in a panel-data setting, rather than recovery speed after a single disruption event.

$$Volatility_{i,t} = sd(Return_{i,m,t}), \quad (5)$$

where $Return_{i,m,t}$ denotes the monthly stock return of firm i in month m of year t . A higher value of $Volatility_{i,t}$ indicates greater financial fluctuation and weaker organizational stability. Therefore, volatility is reverse-coded when constructing the composite resilience index.

Because the medium-term performance dimension and financial stability dimension differ in scale and direction, both dimensions are winsorized at the 1% and 99% levels and then normalized before aggregation. The medium-term performance dimension is normalized within each year as follows:

$$MTPerfNorm_{i,t} = \frac{MTPerf_{i,t} - \min(MTPerf_t)}{\max(MTPerf_t) - \min(MTPerf_t)}, \quad (6)$$

The financial stability dimension is reverse-normalized within each year as follows:

$$StableNorm_{i,t} = \frac{\max(Volatility_t) - Volatility_{i,t}}{\max(Volatility_t) - \min(Volatility_t)}, \quad (7)$$

The organizational resilience index is then constructed as the average of the normalized medium-term performance dimension and the reverse normalized stability dimension:

$$Res_{i,t} = \frac{MTPerfNorm_{i,t} + StableNorm_{i,t}}{2}. \quad (8)$$

A larger value of $Res_{i,t}$ indicates stronger organizational resilience. This construction ensures that organizational resilience reflects both sustained performance capacity and financial stability under uncertainty. It is intended to capture firms' broader absorptive, restorative, and adaptive capacity in a panel-data setting, rather than recovery speed after a single disruption event.

3.2.3. Control Variables

Leverage (*Lev*). A higher leverage ratio increases debt servicing pressure and reduces firms' room for adjustment under external shocks [42], which is unfavorable to organizational resilience. Accordingly, *Lev* is predicted to exert a negative influence on organizational resilience.

Sales growth rate (*Growth*). A higher sales growth rate usually indicates stronger market expansion capacity and also helps strengthen the foundation for risk resistance and operational continuity [2]. *Growth* is therefore expected to have a positive effect on organizational resilience.

Operating cash flow (*CashFlow*). Firms with stronger operating cash flows may be more inclined to maintain short-term operating arrangements rather than invest in long-term capability building [43], which may weaken organizational resilience. *CashFlow* may therefore have a negative effect on organizational resilience.

Board independence (*Indep*). A higher proportion of independent directors can raise coordination burdens in decision-making and slow down corporate response under shock scenarios [44], which is unfavorable to the formation of organizational resilience. *Indep* is therefore expected to have an adverse effect on organizational resilience.

Return on assets (*ROA*). A higher return on assets reflects stronger profitability [45] and also implies that firms possess more abundant resources to maintain operational stability and undertake adaptive adjustment. *ROA* is therefore expected to have a positive effect on organizational resilience.

Executives' digital background (*Del*). Executives with digital backgrounds are more familiar with AI-related concepts and their modes of expression and are also more capable of technological narrative construction and market communication. When substantive investment is insufficient, such executives may be more likely to reinforce AI narrative bias and thereby weaken organizational resilience. *Del* is therefore anticipated to weaken organizational resilience.

Firm age (*Age*). Older firms are usually more likely to accumulate stable resources, richer experience, and more mature risk response mechanisms [46], which enhances their adaptability in complex environments. Based on this logic, *Age* is expected to have a positive effect on organizational resilience.

3.2.4. Mediating Variables

Legitimacy is not directly observable, so this study uses observable organizational outcomes that correspond to the three legitimacy dimensions discussed above. The mediating variables are selected to capture how stakeholders provide resources, evaluate managerial conduct, and interpret the firm's operational reliability. Accordingly, patient capital and net trade credit are used to proxy pragmatic legitimacy, opportunistic executive share reductions are used to proxy moral legitimacy, and supply-chain risk and supply-chain resilience are used to proxy cognitive legitimacy.

Pragmatic legitimacy reflects stakeholders' recognition of a firm's actual value creation capacity. When AI narratives lack support from substantive investment, external stakeholders are more likely to revise downward their assessment of long-term development quality and operating stability. To capture this mechanism, this study uses patient capital and net trade credit as mediating variables for pragmatic legitimacy. These two variables are appropriate observable proxies because pragmatic legitimacy is reflected in whether key external stakeholders are willing to provide stable, relationship-based resources to the firm.

First, patient capital is used to measure the degree of support that long-term investors provide for the firm's development prospects. Long-term institutional investors are more likely to hold shares when they recognize the firm's sustained growth potential and strategic credibility. Therefore, a lower level of patient capital indicates weaker recognition from long-term capital providers and thus weaker pragmatic legitimacy. In this study, patient capital is proxied by long-term institutional investors' ownership share. The specific formula is presented as follows:

$$PC_{it} = \frac{LongInstShare_{it}}{TotalShare_{it}} \quad (9)$$

where the variable captures the patient-capital level of firm *i* in year *t*. *LongInstShare* denotes the shares owned by long-term institutional investors. *TotalShare_{it}* denotes the firm's total outstanding shares. A larger value of this measure indicates that the firm can obtain stronger support from long-term capital. Long-term capital usually places greater weight on firms' sustained growth capacity and can also provide stable resources for coping with external shocks.

Second, net trade credit is used to measure business partners' trust in the firm's operating capacity and contract fulfillment ability. Trade credit is embedded in supplier-customer relationships and reflects whether counterparties are willing to provide financing and transaction support to the firm. It is therefore an observable manifestation of pragmatic

legitimacy in the firm's transaction network. In this study, we operationalize net trade credit as the balance of trade credit received from counterparties after subtracting the trade credit provided by the firm to others. Its specific formula is given as follows:

$$Credit_{it} = \frac{AP_{it} + NP_{it} + Advances_{it} - AR_{it} - NR_{it} - Prepay_{it}}{TA_{it}} \quad (10)$$

where *AP* denotes accounts payable, *NP* denotes notes payable, *Advances* denotes advances from customers, *AR* denotes accounts receivable, *NR* denotes notes receivable, *Prepay* denotes prepayments, and *TA* denotes total assets. A larger value of this measure indicates stronger net trade-credit support received by the firm from suppliers and customers. Net trade credit reflects the firm's credit foundation within the transaction network and also captures external business partners' recognition of its operating stability.

Moral legitimacy reflects stakeholders' assessment of whether corporate behavior is proper and whether governance practices are trustworthy. When firms project an image of technological advancement through AI narratives without support from substantive investment, external stakeholders may interpret such behavior as managerial opportunism. Opportunistic executive share reductions are suitable proxies for moral legitimacy because they capture whether managers appear to use information advantages or market narratives for private benefit. To capture this mechanism, this study uses the amount of opportunistic executive share reduction and the frequency of opportunistic executive share reduction as mediating variables for moral legitimacy.

Following the insider trading classification framework of [47], this study divides executive share reductions into routine reductions and opportunistic reductions. Specifically, we employ an event-study design to estimate abnormal stock returns following each executive share reduction transaction. When a given share reduction predicts a negative stock return over the following month, it is identified as an opportunistic reduction. Otherwise, it is treated as a routine reduction. On this basis, opportunistic reduction transactions are aggregated at the firm-year level, from which two mediator measures are derived. The first is the amount of opportunistic executive share reduction, which is defined as the total amount of all opportunistic reductions by executives in a given year. The second is the frequency of opportunistic executive share reduction, which is defined as the total number of such transactions in a given year. Following prior studies, both variables are transformed using the logarithm. For firm-year observations with no opportunistic executive share reduction, the variable is set to 0.

$$SELLA_{it} = \ln(1 + \sum_{k=1}^{K_{it}} Amount_{ikt}^{opp}) \quad (11)$$

$$SELLT_{it} = \ln(1 + \sum_{k=1}^{K_{it}} N_{ikt}^{opp}) \quad (12)$$

where $SELLA_{it}$ denotes the amount of opportunistic executive share reduction for firm *i* in year *t*, and $SELLT_{it}$ denotes the frequency of opportunistic executive share reduction for firm *i* in year *t*. *Amount* denotes the transaction amount of the *K*-th opportunistic share reduction by firm *i* in year *t*. *N* is an indicator for opportunistic share reduction. It takes the value of 1 if the transaction is identified as opportunistic and 0 otherwise. Higher values of these indicators imply stronger insider opportunism and a greater degree of damage to moral legitimacy.

Cognitive legitimacy reflects the extent to which external stakeholders can understand firms' actual capabilities and strategic logic. When AI narratives lack support from real capabilities, supply-chain partners may find it difficult to assess firms' operational stability.

Supply-chain risk and supply-chain resilience are suitable proxies for cognitive legitimacy because they capture whether the firm's external operating network perceives greater uncertainty and whether the supply-chain system can maintain stable coordination under shocks. To capture this mechanism, this study uses supply-chain risk exposure and supply-chain resilience capacity as mediating variables for cognitive legitimacy.

First, supply-chain risk exposure serves as an indicator of the uncertainty that firms perceive in their supply chains. Following the text-based measurement approach of Wu [48], this study constructs a supply-chain risk indicator from annual report texts. Specifically, when supply-chain terms and risk terms co-occur within a short text window, the text is treated as reflecting the supply-chain risk faced by the firm. The dictionary construction procedure and calculation details are reported in Appendix B. The specific formula is given as follows.

$$SCR_{it} = \frac{1}{T_{it}} \sum_{w=1}^{T_{it}} [1(w \in SC) \times 1(|w - r| \leq 15)] \times 100 \quad (13)$$

Second, supply-chain resilience is used to measure the ability of the firm's supply chain to maintain operation and restore adjustment under external shocks. Stronger supply-chain resilience indicates that the firm's interorganizational relationships are more stable, interpretable, and capable of coordinated adaptation. Drawing on Bruckler, et al. [49], supply-chain resilience is understood as the ability to preserve chain stability under adverse shocks and to recover to a normal state through structural and process adjustment. On this basis, this study uses a composite index to capture supply-chain resilience. The specific formula is given as follows:

$$SCS_{it} = \sum_{k=1}^K \omega_k Z_{kit} \quad (14)$$

where SCS_{it} denotes the supply-chain resilience of firm i in year t . Z_{kit} denotes the k -th standardized indicator of supply-chain resilience. The specific indicator system is reported in the Appendix B. ω_k denotes the corresponding weight. A higher value of this indicator suggests that the firm's supply chain is better able to maintain operation under shocks and to achieve recovery and adjustment more rapidly.

3.3. Model Construction

3.3.1. Double-Debiased Machine Learning

This study adopts double-debiased machine learning (DDML) as the main estimation framework to examine the estimated effect of AI washing on organizational resilience within a complex organizational system. The purpose of using DDML is to estimate the coefficient of interest while allowing the nuisance components associated with firm characteristics and other conditioning variables to be modeled flexibly. AI washing may affect resilience through interdependent relationships among technological narratives, actual AI investment, firm characteristics, governance structures, financial conditions, and external institutional environments. These relationships may be nonlinear and difficult to capture through a purely linear specification. Conventional two-way fixed effects models impose a prespecified functional form on the nuisance components. When multiple organizational factors jointly affect AI washing and organizational resilience, such a restrictive specification may generate functional form bias and reduce estimation reliability.

DDML provides a suitable framework for this study because it decouples the learning of nuisance components from the inference on the coefficient of interest [50]. Specifically, machine-learning algorithms are applied to flexibly learn the conditional expectation of organizational resilience and the conditional mean of AI washing, conditional on the control variables and fixed effects. The residualized outcome and residualized treatment are then used to estimate the coefficient of interest linking AI washing to organizational

resilience. The interpretation of this coefficient relies on the conditional independence assumption that, after conditioning on observed controls and fixed effects, the remaining variation in AI washing is not systematically correlated with unobserved determinants of organizational resilience. By combining Neyman orthogonality and cross-fitting, DDML reduces the sensitivity of the final estimate to regularization bias and overfitting in the nuisance function estimation.

The empirical specification is written as follows:

$$Res_{i,t} = g(X_{i,t}) + \theta AIW_{i,t} + \varepsilon_{i,t}, \quad (15)$$

$$AIW_{i,t} = m(X_{i,t}) + v_{i,t}, \quad (16)$$

where $Res_{i,t}$ denotes organizational resilience, $AIW_{i,t}$ denotes AI washing, and $X_{i,t}$ contains observed firm characteristics and two fixed-effect dimensions, namely firm and year effects. The functions $g(\cdot)$ and $m(\cdot)$ are not restricted to a linear form and are flexibly learned through machine-learning algorithms.

To avoid overfitting, the sample is partitioned into K folds. For each fold, the nuisance functions are estimated using data from the remaining folds, and the fitted functions are then applied to predict organizational resilience and AI washing within the held-out fold. This cross-fitting procedure ensures that the residuals used in the final stage are obtained from out-of-sample predictions.

$$E[(Res_{i,t} - g(X_{i,t}) - \theta(AIW_{i,t} - m(X_{i,t}))) (AIW_{i,t} - m(X_{i,t}))] = 0. \quad (17)$$

The resulting estimate captures the average partial association between AI washing and organizational resilience after flexibly accounting for observed firm-level attributes and fixed-effect structures. Under a conventional conditional exogeneity assumption, the estimate can be viewed as the partial effect of AI washing on organizational resilience. To strengthen the credibility of our empirical analysis, we implement multiple robustness tests, using alternative machine learning learners, alternative fold specifications, alternative variable constructions, propensity score matched samples, expanded fixed effects, and supplementary instrumental variable analysis.

3.3.2. Deep Learning Instrumental Variable Approach

Given the potential endogeneity of AI washing, this study supplements the benchmark DDML specification with a deep learning-based automatic instrumental variable approach. The main endogeneity concerns arise from reverse causality and omitted variables. Firms with lower organizational resilience may have stronger incentives to engage in symbolic technological narratives, while unobserved firm characteristics may simultaneously affect AI washing and organizational resilience. To assess whether the main findings are sensitive to these concerns, we follow Yuan, et al. [51] and use AutoIV to learn instrumental variable representations from a set of theoretically motivated candidate instruments.

Unlike traditional instrumental variable methods, AutoIV does not rely on a single pre-specified instrument. Instead, it uses representation learning to construct latent instrumental variable representations from theoretically motivated candidate variables that are informative for identification. Its central idea is to strengthen the correlation between the learned instrumental representation and the treatment variable, while weakening its direct association with the outcome variable. In this way, the model learns a representation that is relevant to AI washing but less directly predictive of organizational resilience.

Let $\hat{D}_{it}$ denote the predicted treatment variable, and let $\hat{Y}_{it}$ denote the predicted outcome variable. Let Z_{it} denote the candidate-instrument vector. Let X_{it} represent the control-variable vector, which also includes fixed-effect information. The candidate instruments

are constructed before model estimation and capture firms' exposure to external AI-related shocks or AI-oriented policy environments. Specifically, the candidate set includes regional AI-related narrative pressure, pilot-zone exposure interacted with firms' baseline AI foundation, and the post-ChatGPT shock interacted with firms' baseline AI foundation. AutoIV first uses a representation learning network to generate a latent instrumental variable representation from the candidate instruments:

$$\tilde{Z}_{it} = f_{\omega}(Z_{it}, X_{it}), \quad (18)$$

where $f_{\omega}(\cdot)$ denotes the instrumental variable representation learning network. This network is used to extract exogenous information from the candidate variables that is relevant to the treatment variable. At the same time, the model imposes additional constraints to reduce the direct explanatory power of Z_{it} for the outcome variable, thereby mitigating the risk of violating the exclusion restriction. In the implementation, this restriction is encouraged through an adversarial outcome-prediction component based on a gradient-reversal layer.

At the first-stage step, AI washing is predicted using this latent instrumental representation together with the control variables:

$$\hat{D}_{it} = h_{\theta}(\tilde{Z}_{it}, X_{it}), \quad (19)$$

where $h_{\theta}(\cdot)$ denotes the first-stage treatment prediction network. $\hat{D}_{it}$ represents the instrument-induced predicted component of AI washing. This stage is designed to extract the component of the treatment variable that is closer to exogenous variation. Firm fixed-effect indicators and year fixed-effect indicators are included in X_{it} , together with the firm-level control variables used in the baseline specification.

At the second-stage step, the first-stage fitted treatment and control variables are jointly entered into the outcome network:

$$\hat{Y}_{it} = g_{\phi}(\hat{D}_{it}, X_{it}), \quad (20)$$

where $g_{\phi}(\cdot)$ denotes the second-stage outcome network. This specification allows the treatment effect on the outcome to be modeled nonlinearly. It also avoids the strict functional-form restrictions imposed by conventional linear instrumental variable models.

Further, this study calculates the treatment's average marginal effect on the outcome through numerical perturbation. Specifically, holding the control variables constant, a small perturbation δ is imposed on $\hat{D}$, and the average change in the model-predicted outcome is computed as follows:

$$AME = \frac{1}{N} \sum_{i=1}^N \frac{g_{\phi}(\hat{D}_i + \delta, X_i) - g_{\phi}(\hat{D}_i, X_i)}{\delta}, \quad (21)$$

where N denotes the sample size, and AME denotes the mean marginal response of the outcome to changes in the treatment. This indicator captures the net outcome variation associated with a marginal change in the treatment variable within the deep instrumental variable framework. Statistical inference is based on firm-level cluster bootstrap resampling. To enhance the stability and generalizability of the learned representation, the sample is divided into training and testing sets with a fixed random seed, and model performance is evaluated using out-of-sample prediction.

Compared with conventional instrumental variable methods, AutoIV can learn instrumental representations from multiple candidate variables and accommodate nonlinear

relationships between instruments, treatment, controls, and outcome. This feature is useful in the present setting, where AI washing may be shaped by several external AI-related shocks and policy environments rather than by a single source of variation. The identifying interpretation follows the standard instrumental-variable logic: conditional on firm characteristics and fixed effects, the candidate instruments are expected to affect organizational resilience through their impact on AI washing. Under this exclusion restriction, the AutoIV results provide additional evidence that the negative association between AI washing and organizational resilience is not driven by the baseline functional-form choice or by conventional observable confounding.

3.4. Descriptive Statistics and Variable Distributions

Table 1 presents summary statistics for the key variables. The final sample comprises 35,764 firm-year records. The mean of AI washing is -3.38×10^{-5} and the median is 8.86×10^{-5} , which suggests that firms' AI narratives and actual AI investment are broadly aligned on average. Its standard deviation is 0.003, while the minimum and maximum values are -0.006 and 0.007 , respectively. This indicates considerable variation across firms in narrative deviation. Organizational resilience has a mean value of 0.867 and a median value of 0.884, which implies that firms generally exhibit a high level of resilience. The minimum value, however, is 0.128, which suggests that some firms still have a weak resilience base. Figure 2 shows the temporal and cross-industry distribution of AI washing, whereas Figure 3 presents the corresponding distribution of AI investment. The relative level of AI investment rises overall after 2019, yet some industries still display a pattern in which narrative expansion precedes actual investment. Figure 4 shows that organizational resilience remains at a relatively high level from 2019 to 2022 but declines somewhat in 2023 and 2024. This suggests that the post-pandemic recovery of corporate resilience did not continue to strengthen. Overall, the sample data capture variation in firms' AI narrative deviation and organizational resilience well and provide a reliable empirical foundation for the following analysis.

Table 1. Descriptive statistics.

Variable	N	Mean	SD	Min	Median	Max
<i>AIW</i>	35,764	-3.38×10^{-5}	0.003	-0.006	8.86×10^{-5}	0.007
<i>Res</i>	35,764	0.867	0.075	0.128	0.884	0.974
<i>Lev</i>	35,764	0.402	0.198	0.028	0.394	0.908
<i>CashFlow</i>	35,764	0.050	0.067	-0.222	0.049	0.267
<i>Growth</i>	35,764	0.146	0.368	-0.673	0.083	5.076
<i>Indep</i>	35,764	0.377	0.0537	0.286	0.364	0.600
<i>ROA</i>	35,764	0.042	0.0596	-0.556	0.041	0.222
<i>Age</i>	35,764	2.929	0.349	0.000	2.944	4.290
<i>Del</i>	35,764	0.146	0.353	0	0	1

Note: This table presents summary statistics for the key variables. *AIW* captures the extent of AI washing and is calculated as the difference between a firm's within-industry percentile position in AI disclosure and its corresponding percentile position in actual AI investment. *Res* denotes organizational resilience. *Lev* is the ratio of total liabilities to total assets. *CashFlow* is operating cash flow scaled by total assets. *Growth* measures the operating revenue growth rate. *Indep* captures the share of independent directors on the board. *ROA* represents return on assets. *Age* denotes firm age. *Del* is an indicator variable for whether corporate executives have a digital background.

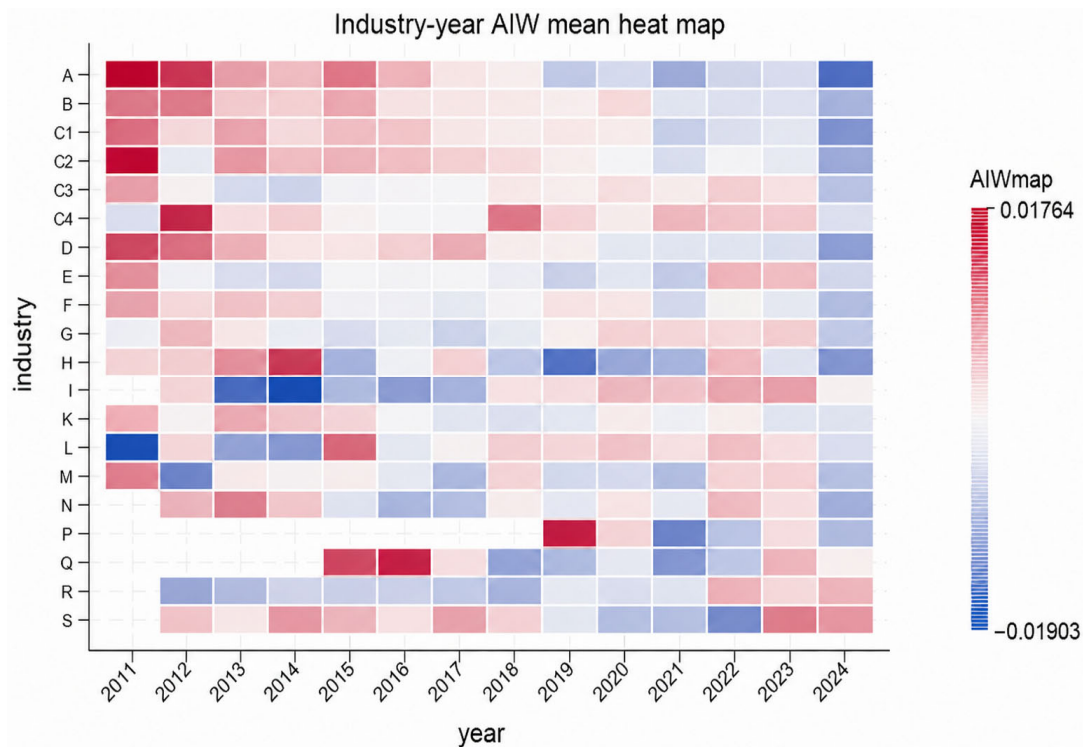


Figure 2. Heat map of the relative level of AI washing by industry and year. Note: Redder shading indicates a relatively higher level of AI washing in a given industry-year, whereas bluer shading indicates a relatively lower level. The results show clear cross-industry differences in AI washing. They also show stage-specific fluctuations over time rather than a monotonic upward trend.

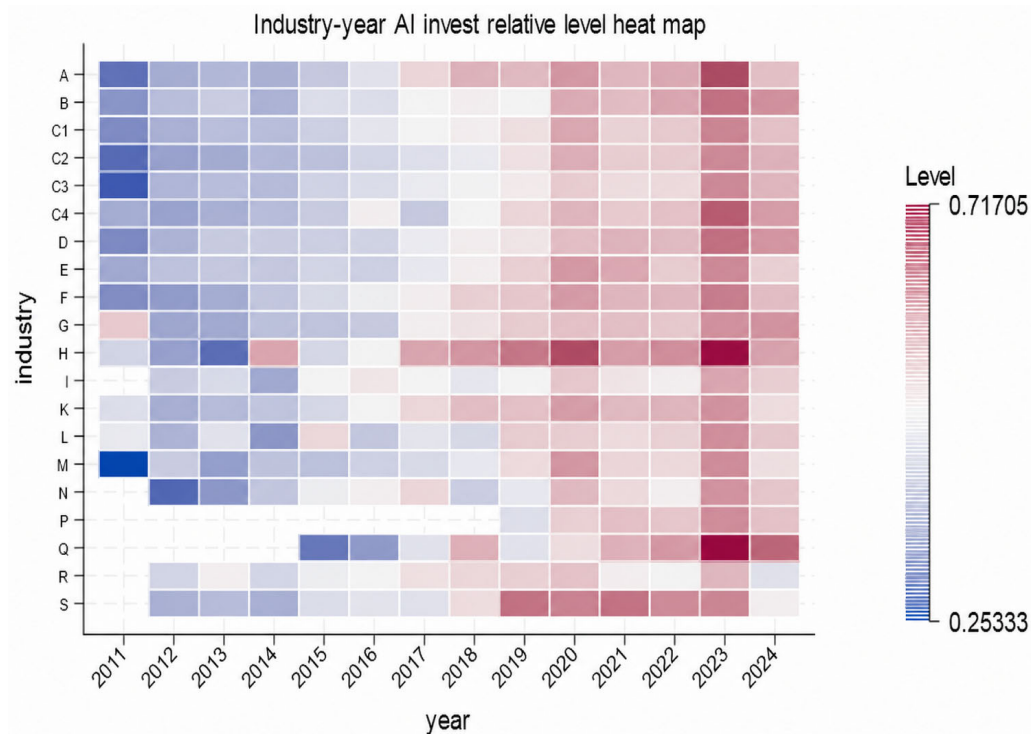


Figure 3. Heat map of the relative level of AI investment by industry and year. Note: Redder shading indicates a relatively higher level of AI investment in a given industry-year, whereas bluer shading indicates a relatively lower level. The results show clear cross-industry differences in AI investment. They also show stage-specific fluctuations over time rather than a monotonic upward trend.

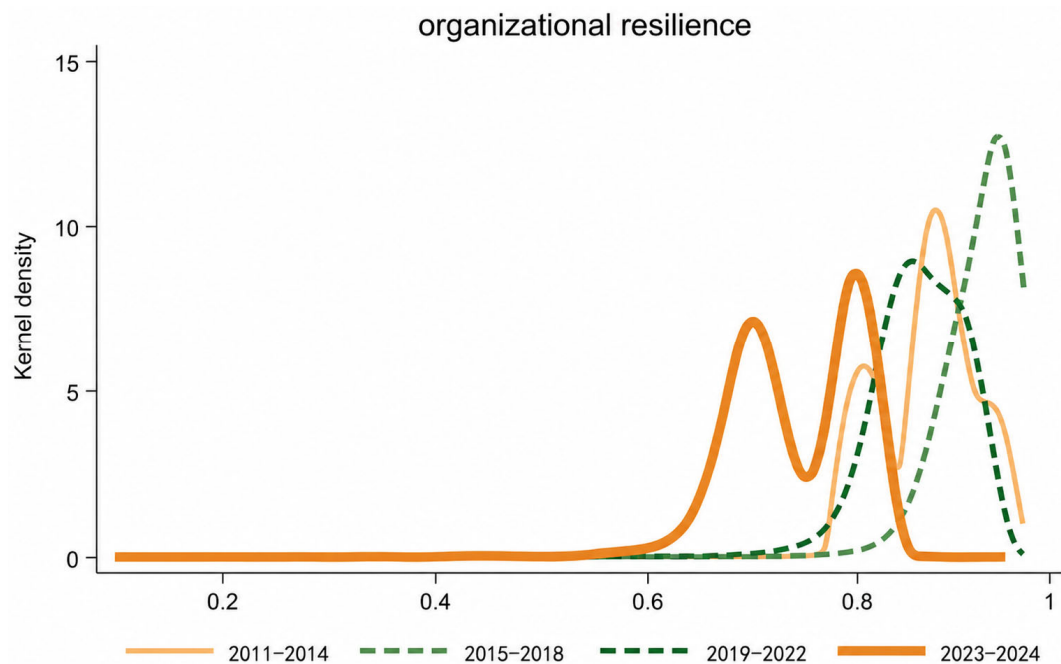


Figure 4. Temporal distribution of organizational resilience. Note: This figure presents the kernel density distributions of organizational resilience across four periods: 2011–2014, 2015–2018, 2019–2022, and 2023–2024. The horizontal axis represents the level of organizational resilience, and the vertical axis reports kernel density. The distributional shifts across periods illustrate the temporal evolution of firms' organizational resilience.

4. Empirical Results

4.1. Baseline Estimation Results

Table 2 summarizes the baseline estimates for the relationship between AI washing and organizational resilience. Column (1) of Table 2 reports the baseline result. The coefficient on *AIW* is negative and statistically significant at the 1% level, suggesting that AI washing is associated with lower organizational resilience under the maintained DDML assumptions. We further scale this effect by the absolute value of the mean annual rate of change in organizational resilience, which is 0.025. The estimated effect is equivalent to approximately 3.276% ($0.273 \times 0.003 / 0.025$) of the average annual change in organizational resilience. These results indicate that the estimated relationship between AI washing and organizational resilience is not only statistically significant but also economically meaningful.

Column (2) presents the results from a conventional two-way fixed-effects specification. The estimated coefficient of AI washing is -0.231 and remains significant at the 1% level. This indicates that the negative estimated relationship between AI washing and organizational resilience also appears under the traditional fixed-effects framework. Although the coefficient is smaller than the DDML estimate, its sign and significance remain unchanged. This pattern suggests that the core conclusion is not tied to any single estimation strategy.

Following Chen, et al. [52], we additionally include industry-year fixed effects and province-year fixed effects. Columns (3) and (4) present the results after adding these fixed effects. In both columns, the estimated coefficient on AI washing equals -0.273 and remains statistically significant at the 1 percent level. This indicates that the negative estimated relationship between AI washing and organizational resilience persists after accounting for time-varying industry- and region-level shocks.

Table 2. Results of baseline regression.

	(1) <i>Res</i>	(2) <i>Res</i>	(3) <i>Res</i>	(4) <i>Res</i>
<i>AIW</i>	−0.273 *** (0.055)	−0.231 *** (0.079)	−0.273 *** (0.059)	−0.273 *** (0.057)
<i>Constant</i>	−0.000 ** (0.000)	0.840 *** (0.007)	−0.000 ** (0.000)	−0.000 (0.001)
Controls	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Industry × Year FE	×	×	✓	×
Province × Year FE	×	×	×	✓
<i>N</i>	35,764	35,764	35,764	35,764

Note: This table presents the baseline regression estimates for the relationship between AI washing and organizational resilience. *Res* denotes organizational resilience. *AIW* represents AI washing. Columns (1) and (2) report the DDML and OLS estimation results, respectively. Columns (3) and (4) further control for industry and year fixed effects and province and year fixed effects, respectively. Firm and year fixed effects are included as indicated. ✓ indicates that the corresponding controls or fixed effects are included, whereas × indicates that they are not included. Standard errors are reported in parentheses. *** and ** denote significance at the 1% and 5% levels, respectively.

4.2. Endogeneity Analysis Based on Deep Learning Instrumental Variables

Potential endogeneity may arise from reverse causality and omitted variables. Firms with weaker organizational resilience may have greater incentives to undertake image management activities and intensify AI-related narratives, while unobserved firm characteristics may simultaneously affect AI washing and organizational resilience. To address such endogeneity concerns, we further employ a deep instrumental-variable strategy based on AutoIV.

We construct three groups of candidate instruments. The first is the average intensity of AI narratives for peer firms located in the same region, which reflects local pressure surrounding AI narratives while excluding the focal firm. The second is the interaction between the pilot zone policy shock and firms' pre-policy AI textual intensity, which captures policy-induced variation in AI narrative incentives. The third is the interaction between the ChatGPT shock and firms' AI textual intensity before the emergence of GPT, which captures external technological attention shocks. These candidate instruments are expected to affect firms' AI washing through external narrative pressure, policy incentives, and general technological attention rather than through firms' internal resilience conditions.

Following Yuan, Wu, Kuang, Li, Wu, Wu and Lin [51], AutoIV is used to learn an instrumental variable representation from the above candidate instruments. The method is designed to strengthen the association between the learned instrument representation and AI washing while reducing its direct dependence on organizational resilience. It therefore allows the instrumental variable structure to be learned from multiple theoretically motivated candidate variables rather than relying on a single pre-specified instrument. In addition, the two-stage neural network structure makes it possible to capture nonlinear relationships between AI washing and organizational resilience. The average marginal effect of AI washing is then calculated through numerical perturbation in the second-stage outcome model.

Table 3A presents the AutoIV estimation results. Panel A reports the model-performance statistics. The first-stage treatment-prediction fit indicates that the learned instrumental representation has strong in-sample predictive power for AI washing. Because the reported first-stage R^2 is calculated from the training sample, it reflects the in-sample fit of the treatment-prediction step rather than the validity of the learned instrument representation. The outcome-prediction stage also shows satisfactory predictive performance,

suggesting that the trained AutoIV framework captures meaningful variation in organizational resilience. More importantly, the estimated average marginal effect of AI washing is -0.447 and statistically significant at the 1% level. This result provides supplementary deep instrumental-variable evidence that AI washing is negatively associated with organizational resilience under the AutoIV-based framework. The identifying interpretation continues to rely on the maintained exclusion restriction that the candidate variables affect organizational resilience mainly through AI washing, conditional on firm characteristics and fixed effects. Therefore, the AutoIV results provide additional support for the baseline negative association under the maintained IV assumptions.

Table 3. AutoIV for the effect of AI washing on organizational resilience. (A) AutoIV estimates; (B) linear IV benchmark estimates.

(A) AutoIV	
Variable	Res (1)
<i>AIW Average Marginal Effect</i>	-0.447^{***}
<i>First-stage R^2</i>	0.997
<i>Second-stage R^2</i>	0.829
(B) Linear IV	
<i>AIW</i>	-6.160^{***}
Sanderson-Windmeijer F test	119.859
Kleibergen-Paap rk LM	103.817
Controls	✓
Firm FE	✓
Year FE	✓
N	35,764

Note: This table presents the AutoIV-based analysis for addressing potential endogeneity. Panel A presents the AutoIV estimates. The reported coefficient is the *AME* of *AIW* on *Res* derived from the second-stage outcome model. The first-stage R^2 refers to the treatment-prediction fit evaluated in the training sample. Significance levels for the AutoIV average marginal effect are based on firm-level cluster bootstrap inference. Panel B reports the linear IV benchmark estimates using the instrument constructed from the AutoIV procedure. All regressions include the full set of control variables, firm fixed effects, and year fixed effects. The Sanderson–Windmeijer F-statistic and Kleibergen–Paap rk LM statistic are provided, where applicable, to evaluate instrument relevance and underidentification. ✓ indicates that the corresponding controls or fixed effects are included. *** denotes statistical significance at the 1% level.

Table 3B reports the linear IV benchmark based on the instrument constructed from the AutoIV procedure. The estimated coefficient of *AIW* remains negative and statistically significant. The Sanderson–Windmeijer F statistic is 119.859, and the Kleibergen–Paap rk LM statistic equals 103.817, indicating that the learned instrument representation has sufficient relevance in the linear IV benchmark. Overall, the deep instrumental variable analysis reinforces the baseline conclusion that AI washing weakens organizational resilience. Additional details on the construction of the AutoIV candidate variables and robustness checks based on alternative candidate variable sets are reported in Appendix D.

4.3. Robustness Tests

To validate the robustness of the empirical research, we further conduct a series of additional robustness checks employing multiple methodological approaches.

4.3.1. Alternative AI Washing Measures

To ensure that the measurement of the core variables does not drive the baseline results, this study also conducts tests with alternative explanatory variables. On the one hand, the indicator of firms' actual AI input in the AI washing measure is replaced with the number of granted AI patents [3]. On the other hand, the indicator of AI disclosure intensity

is replaced. Specifically, the baseline measure based on large language model identification is replaced with a word frequency-based measure of AI disclosure [37], and three new AI washing indicators are reconstructed accordingly. As reported in Columns (1)–(3) of Table 4, after remeasuring the key variables, the estimated coefficients on AI washing are directionally consistent, and their statistical significance does not change materially. These findings further support the reliability of the main conclusions.

Table 4. Robustness checks using alternative variable constructions.

	Alternative Explanatory Variable			Alternative Dependent Variable
	(1) <i>Res</i>	(2) <i>Res</i>	(3) <i>Res</i>	(4) <i>FPG</i>
<i>AIW</i>				−92.329 ** (46.530)
<i>AIW2</i>	−0.361 *** (0.069)			
<i>AIW3</i>		−0.311 *** (0.055)		
<i>AIW4</i>			−0.381 *** (0.063)	
<i>Constant</i>	−0.000 ** (0.000)	−0.000 (0.001)	−0.000 (0.000)	−0.000 (0.000)
<i>N</i>	35,764	35,764	35,764	35,764

Note: This table presents robustness checks using alternative variable constructions. Columns (1)–(3) replace the baseline AI washing measure with alternative *AIW* indicators, while Column (4) replaces the dependent variable with *FPG*. Standard errors are reported in parentheses. *** and ** denote statistical significance at the 1% and 5% levels, respectively.

4.3.2. Alternative Dependent Variable

To strengthen the robustness of the findings, this study follows Xu, et al. [53] and further replaces the dependent variable with the dispersion of employee income per capita. Although this indicator is not a direct measure of organizational resilience, it provides an operationally relevant proxy for the internal coordination foundation of resilience. It can reflect the internal foundations of firms' shock response from the perspective of income distribution and organizational coordination. Excessive income disparity among employees often signals a weaker sense of fairness, lower cooperative efficiency, and reduced cohesion, which may in turn undermine firms' operational continuity and risk response capacity. As reported in Column (4) of Table 4, the estimated coefficient on AI washing remains statistically significant after replacing the dependent variable. This finding suggests that the main conclusion regarding the resilience effect of AI washing is robust.

4.3.3. Alternative Clustering Levels

To rule out the possibility that the baseline evidence is driven by the choice of standard-error clustering, this study further conducts a clustering replacement test. Specifically, while keeping the model specification unchanged, standard errors are clustered under four alternative clustering dimensions, namely firm, industry, city, and province, respectively. According to the results in Table 5, the estimated coefficient on AI washing remains directionally consistent under all clustering schemes, and its statistical significance does not change materially. This indicates that the main results are not sensitive to the selected clustering level and are therefore robust.

Table 5. Robustness checks using alternative clustering levels.

	Alternative Clustering Levels			
	(1) <i>Res</i>	(2) <i>Res</i>	(3) <i>Res</i>	(4) <i>Res</i>
<i>AIW</i>	−0.273 *** (0.059)	−0.273 *** (0.057)	−0.273 *** (0.066)	−0.273 *** (0.076)
<i>Constant</i>	0.027 ** (0.003)	0.022 (0.003)	0.017 *** (0.002)	0.023 *** (0.002)
Controls	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
<i>N</i>	35,764	35,764	35,764	35,764

Note: Columns (1)–(4) report robustness tests with standard errors clustered at the firm, industry, city, and province levels, respectively. The model specification remains unchanged across columns. ✓ indicates that the corresponding controls or fixed effects are included. Standard errors are reported in parentheses. *** and ** denote statistical significance at the 1% and 5% levels, respectively.

4.3.4. Alternative Machine Learning Methods

Following Shi, et al. [54], this study further conducts a replacement test using alternative machine learning methods. While keeping the sample scope, variable definitions, and control variable specification unchanged, the model is re-estimated using Lasso, support vector machine, random forest, and OLS. The results reported in columns (1)–(4) of Table 6 show that the estimated AI washing effect remains consistent in direction across different algorithms, and the conclusions do not change materially. This indicates that the estimated negative association between AI washing and organizational resilience is robust across alternative machine-learning methods.

Table 6. Robustness checks with alternative learners and cross-fitting folds specification.

	Alternative Learners				Alternative Cross-Fitting Folds	
	(1) <i>Res</i>	(2) <i>Res</i>	(3) <i>Res</i>	(4) <i>Res</i>	(5) <i>Res</i>	(6) <i>Res</i>
<i>AIW</i>	−0.242 *** (0.075)	−0.461 *** (0.054)	−0.346 *** (0.055)	−0.213 *** (0.075)	−0.288 *** (0.055)	−0.278 *** (0.055)
<i>Constant</i>	0.000 (0.000)	0.052 *** (0.000)	−0.005 *** (0.000)	−0.000 (0.000)	−0.000 ** (0.000)	−0.000 ** (0.000)
learners	LassoCv	SVM	RF	OLS	XGB	XGB
cross-fitting folds	5	5	5	5	8	10
Controls	✓	✓	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
<i>N</i>	35,764	35,764	35,764	35,764	35,764	35,764

Note: Columns (1)–(4) replace the baseline learner with four alternative estimators: Lasso, SVM, random forest, and OLS. Columns (5)–(6) re-estimate the model under different cross-fitting partition schemes. All reported specifications include the indicated control variables, firm fixed effects, and year fixed effects. ✓ indicates that the corresponding controls or fixed effects are included. Values in parentheses are standard errors. *** and ** indicate rejection of the null hypothesis at the 1% and 5% levels, respectively.

4.3.5. Alternative Machine Cross-Fitting Folds

Given that cross-fitting performance may vary with the selected number of folds and may affect finite sample-based estimates in double debiased machine learning [55], we further examine robustness by altering the fold setting. In addition to the baseline setting, we set the cross-fitting folds to 8 and 10, respectively. The results in columns (5) and (6) of Table 6 show that, after the fold specification is changed, the estimated coefficient on

AI washing remains directionally consistent and its significance level does not change materially. This indicates that our main findings are not affected by the chosen number of cross-fitting folds and are therefore highly robust.

4.3.6. Propensity Score Matching

Firms with different levels of AI washing may differ systematically in observable characteristics, which could affect the inference regarding the relationship between AI washing and organizational resilience. To mitigate this concern, we use propensity score matching (PSM) to construct a comparable matched sample between firms with higher and lower AI washing. Specifically, firms with AI washing indicators above the sample median are defined as the treatment group, while firms at or below the sample median serve as the initial control group. The matching covariates are the same firm-level control variables used in the baseline analysis. As reported in Table 7, the covariate balance test shows that the matching procedure substantially reduces observable differences between the treatment and control groups. After matching, the absolute standardized bias of all covariates is below 3%, indicating that the matched sample achieves satisfactory covariate balance.

Table 7. Covariate balance test after propensity score matching.

Variable	Sample	Treated	Control	%Bias
<i>Lev</i>	U	0.369	0.434	−33.104
	M	0.369	0.370	−0.190
<i>Growth</i>	U	0.135	0.157	−5.860
	M	0.135	0.137	−0.541
<i>CashFlow</i>	U	0.046	0.054	−11.728
	M	0.046	0.046	−0.902
<i>Indep</i>	U	0.377	0.376	1.673
	M	0.377	0.376	0.907
<i>ROA</i>	U	0.042	0.040	−0.586
	M	0.042	0.040	2.156
<i>Del</i>	U	0.153	0.139	3.800
	M	0.153	0.152	0.143
<i>Age</i>	U	2.910	2.947	−10.746
	M	2.910	2.903	1.114

Note: This table reports covariate balance before and after propensity score matching. U refers to observations before matching, whereas M refers to observations retained after matching. The treatment group consists of firms with above-threshold AI washing, while the control group consists of firms with lower AI washing intensity. Standardized bias measures the difference in covariate means between the treatment and control groups, scaled by the pooled standard deviation. Smaller absolute values indicate better covariate balance. After matching, the absolute standardized bias of all covariates is below 3%, suggesting that the matched sample achieves satisfactory covariate balance.

Subsequently, following the matching procedure outlined by Tang, et al. [56], we implement nearest-neighbor matching and pair each treated firm with up to two control firms. The caliper is set at 0.03 to restrict the maximum allowable difference in propensity scores. To examine whether the matching results are sensitive to the matching design, we further adopt two alternative matching strategies. First, we adjust the matching ratio to 1:3 while retaining the caliper restriction, thereby improving matching efficiency and reducing estimation variance. Second, we implement radius matching, which includes all eligible control firms within the caliper and relaxes the restriction on a fixed number of matched controls. The corresponding PSM results are reported in Table 8.

Table 8. Robustness checks with PSM samples.

	(1) <i>Res</i>	PSM (2) <i>Res</i>	(3) <i>Res</i>
<i>AIW</i>	−0.274 *** (0.063)	−0.268 *** (0.059)	−0.306 *** (0.055)
<i>Constant</i>	−0.000 ** (0.000)	−0.000 * (0.000)	−0.000 ** (0.000)
Controls	✓	✓	✓
Firm FE	✓	✓	✓
Year FE	✓	✓	✓
<i>N</i>	27,396	31,421	35,756

Note: Columns (1)–(3) use propensity-score-matched samples constructed through nearest-neighbor matching, 1:3 nearest-neighbor matching, and radius matching, respectively. All reported specifications include the indicated control variables, firm fixed effects, and year fixed effects. ✓ indicates that the corresponding controls or fixed effects are included. Values in parentheses are standard errors. ***, **, and * indicate rejection of the null hypothesis at the 1%, 5%, and 10% levels, respectively.

4.3.7. Additional Controls for Governance and Managerial Characteristics

A potential concern is that firms with higher *AIW* may already differ from other firms in governance quality or managerial characteristics, and that these pre-existing differences may confound the estimated relationship between AI washing and organizational resilience. To address this concern, we further include two additional firm-level controls: internal control quality and managerial overconfidence. These variables capture governance infrastructure and managerial behavioral tendencies beyond the baseline controls and fixed effects already included in the model.

Internal control quality is measured using the DIB Internal Control Index, which is widely used to assess the quality of internal control systems among Chinese listed firms. A higher value indicates stronger internal control quality and better governance infrastructure. This measurement is consistent with China-based accounting and governance research that uses a COSO-based internal control index to capture firms' internal control quality [57]. Managerial overconfidence is measured using a relative compensation concentration approach. Specifically, we calculate the ratio of the compensation of the top three executives to the total compensation of all managers. If this ratio is above the sample mean, the firm-year observation is coded as having managerial overconfidence; otherwise, it is coded as non-overconfident. This proxy captures the extent to which compensation is concentrated among the most influential executives, reflecting managerial dominance and confidence in strategic judgment. It is also consistent with recent research linking managerial or board overconfidence to major strategic decisions and risk-taking behavior [58].

Table 9 reports the results. Column (1) adds internal control quality, Column (2) adds managerial overconfidence, and Column (3) includes both variables simultaneously. The coefficient on *AIW* remains negative and statistically significant across all three specifications. Specifically, the estimated coefficient is −0.363 when internal control quality is added, −0.347 when managerial overconfidence is added, and −0.350 when both variables are included. These results indicate that the negative estimated relationship between AI washing and organizational resilience is not fully explained by observable differences in internal control quality or managerial overconfidence.

Table 9. Addressing potential confounding from governance and managerial characteristics.

	(1) <i>Res</i>	(2) <i>Res</i>	(3) <i>Res</i>
<i>AIW</i>	−0.363 *** (0.059)	−0.347 *** (0.059)	−0.350 *** (0.059)
<i>Cons</i>	−0.001 *** (0.000)	−0.001 *** (0.000)	−0.001 *** (0.000)
Baseline controls	✓	✓	✓
Internal control	✓	×	✓
Managerial overconfidence	×	✓	✓
Firm FE	✓	✓	✓
Year FE	✓	✓	✓
<i>N</i>	31,564	31,513	31,511

Note: This table reports additional robustness checks addressing potential confounding from governance quality and managerial characteristics. Column (1) adds internal control quality, Column (2) adds managerial overconfidence, and Column (3) includes both variables simultaneously. Internal control quality captures firms' governance quality, while managerial overconfidence captures managerial behavioral and strategic tendencies. The baseline controls include leverage, sales growth, operating cash flow, return on assets, firm age, board independence, and executives' digital background. ✓ indicates that the corresponding variable, controls, or fixed effects are included, whereas × indicates that they are not included. All specifications include firm and year fixed effects. Standard errors are reported in parentheses. *** denotes significance at the 1% level.

4.4. Potential Channels

Building on the baseline results, we further investigate the transmission channels through which AI washing affects organizational resilience. According to the theoretical analysis, AI washing may reduce organizational resilience by eroding organizational legitimacy. Specifically, this paper examines three dimensions of legitimacy within the organizational system: pragmatic legitimacy, moral legitimacy, and cognitive legitimacy. To strengthen the temporal ordering of the mediation analysis, this paper sets AI washing as the explanatory variable with a one-period lag, the mediator as a contemporaneous variable, and organizational resilience as the one-period forward outcome. This treatment helps clarify the dynamic sequence in which AI washing first affects organizational legitimacy and then influences future organizational resilience. It also helps reduce simultaneity bias and reverse causality concerns, thereby providing a more rigorous basis for identifying the proposed transmission channels.

This paper first conducts path regressions to examine whether AI washing significantly affects legitimacy-related mediators. The model is specified as follows:

$$M_{it} = \alpha_0 + \alpha_1 AIW_{i,t-1} + \gamma Controls_{i,t-1} + \mu_i + \lambda_t + \varepsilon_{it}, \quad (22)$$

where M_{it} denotes the mediator associated with a specific legitimacy channel. $AIW_{i,t-1}$ denotes AI washing with a one-period lag. $Controls_{i,t-1}$ denotes the lagged control-variable vector. μ_i and λ_t represent firm and year fixed effects, respectively. If α_1 is statistically significant, AI washing is considered to have a significant effect on the corresponding mediator, thereby providing preliminary evidence for the proposed transmission path.

On this basis, this paper further uses CMAverse to conduct causal mediation analysis. Unlike the traditional stepwise regression approach, this paper does not judge the existence of mediation by adding the mediator to the outcome equation and comparing changes in the coefficient of the explanatory variable. Instead, CMAverse directly decomposes the natural direct and indirect components of the effect within a counterfactual framework. The mediator model and outcome model specified in CMAverse are as follows:

$$M_{it} = \theta_0 + \theta_1 AIW_{i,t-1} + \theta_2 Controls_{i,t-1} + \mu_i + \lambda_t + v_{it}, \quad (23)$$

$$Res_{i,t+1} = \delta_0 + \delta_1 AIW_{i,t-1} + \delta_2 M_{it} + \delta_3 Controls_{i,t-1} + \mu_i + \lambda_t + \xi_{i,t+1}. \quad (24)$$

where $Res_{i,t+1}$ denotes the organizational resilience of firm i for period $t + 1$. The first specification serves as the mediator model, which captures the effect of lagged AI washing on the current mediator. The second equation is the outcome model, which captures the effects of lagged AI washing and the current mediator on future organizational resilience. Using these two models, CMAverse partitions the overall effect of AI washing on organizational resilience into its natural direct and natural indirect components as follows:

$$TE = NDE + NIE, \quad (25)$$

where TE represents the overall effect, NDE represents the natural direct component, and NIE represents the natural indirect component transmitted through the mediator. The mediation proportion is further defined as

$$MP = \frac{NIE}{TE}. \quad (26)$$

In the empirical analysis, if α_1 in the path regression is statistically significant, AI washing is considered to significantly affect the proposed mediator. If the natural indirect effect NIE estimated by CMAverse is statistically significant, the corresponding mediator is regarded as an effective channel through which AI washing affects organizational resilience. Therefore, path regression provides preliminary evidence on the direction of the mechanism, whereas CMAverse serves as the main basis for determining whether the mediation effect holds.

The following analysis, therefore, examines whether AI washing is associated with changes in legitimacy-related organizational outcomes and whether these changes help explain the estimated negative association between AI washing and organizational resilience. Patient capital, net trade credit, opportunistic executive share reductions, supply-chain risk, and supply-chain resilience capture observable manifestations of pragmatic, moral, and cognitive legitimacy in firms' financial, governance, and supply-chain relationships. Although these variables do not directly measure stakeholders' subjective legitimacy perceptions, the counterfactual mediation estimates provide model-based evidence consistent with the proposed channels under the maintained mediation assumptions.

4.4.1. Pragmatic Legitimacy

Pragmatic legitimacy emphasizes stakeholders' recognition of a firm's actual capacity to create value and reflects the resource linkages through which the firm is connected to its broader organizational system. This study captures the pragmatic legitimacy mechanism from the perspective of external resource support. Specifically, patient capital and net trade credit are selected as mediating variables. Patient capital reflects long-term investors' support for the firm's development prospects. Net trade credit reflects the trust of business partners, such as suppliers and customers, in the firm's operating capacity.

The regression results of Table 10 show that AI washing is negatively associated with the observable proxies for pragmatic legitimacy. The two-step results suggest that the estimated coefficient of AI washing on patient capital is -696.426 and reaches statistical significance at the 1 percent level. The estimated coefficient of AI washing on net trade credit is -1.625 and remains statistically significant at the 1 percent level. As shown in Columns (1) and (2) of Table 11, the CMAverse estimates provide model-based evidence consistent with a pragmatic-legitimacy channel. The indirect effect through patient capital equals -0.038 and reaches statistical significance at the 1 percent level. The related mediation proportion is 11.475%. The indirect effect through net trade credit equals -0.008 and

reaches statistical significance at the 1 percent level. Its mediation proportion is 2.298%. These findings indicate that patient capital and net trade credit provide evidence consistent with transmission channels linking AI washing to organizational resilience under the maintained mediation assumptions. In addition, the E values for the patient capital path and the net trade credit path are 6.161 and 1.873, respectively. This suggests that the above mediating effects are difficult to attribute entirely to potential unobserved confounders.

Table 10. Mechanism analysis through organizational legitimacy channels.

	(1) <i>PC</i>	(2) <i>NTC</i>	(3) <i>SELLA</i>	(4) <i>SELLT</i>	(5) <i>SCR</i>	(6) <i>SCS</i>
<i>AIW</i>	−696.426 *** (38.789)	−1.625 *** (0.175)	20.016 * (10.854)	2.455 ** (1.194)	1.390 *** (0.293)	−17.310 *** (1.113)
<i>Constant</i>	−0.127 (0.112)	−0.000 (0.001)	0.039 (0.032)	0.005 (0.003)	0.002 *** (0.001)	0.007 ** (0.003)
Controls	✓	✓	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
<i>N</i>	35,764	35,764	35,534	35,534	35,764	35,764

Note: This table presents the path-regression evidence for the proposed organizational legitimacy channels. *PC* refers to patient capital, and *NTC* refers to net trade credit. *SELLA* captures the monetary scale of opportunistic share reductions by executives, while *SELLT* captures the corresponding transaction frequency. *SCR* measures supply-chain risk, and *SCS* measures supply-chain resilience. *AIW* represents AI washing. Columns (1)–(2) test the pragmatic-legitimacy pathway. Columns (3)–(4) test the moral-legitimacy pathway. Columns (5)–(6) test the cognitive-legitimacy pathway. All specifications control for the complete set of covariates, together with firm and year fixed effects. ✓ indicates that the corresponding controls or fixed effects are included. Standard errors are shown in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 11. Causal mediation analysis through organizational legitimacy channels.

	(1) <i>PC</i>	(2) <i>NTC</i>	(3) <i>SELLA</i>	(4) <i>SELLT</i>	(5) <i>SCR</i>	(6) <i>SCS</i>
<i>TE</i>	−0.332 ***	−0.336 ***	−0.353 ***	−0.353 ***	−0.335 ***	−0.337 ***
<i>NDE</i>	−0.294 ***	−0.328 ***	−0.348 ***	−0.349 ***	−0.328 ***	−0.316 ***
<i>NIE</i>	−0.038 ***	−0.008 ***	−0.005 ***	−0.004 ***	−0.007 ***	−0.021 ***
<i>PM</i>	11.475%	2.298%	1.527%	1.249%	2.129%	6.185%
<i>E-Value</i>	6.161	1.873	1.698	1.571	1.820	3.313
Controls	✓	✓	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓

Note: This table documents the counterfactual mediation decomposition used to identify the organizational-legitimacy transmission channels. *PC* denotes patient capital. *NTC* denotes net trade credit. *SELLA* denotes the amount of opportunistic executive share reduction. *SELLT* denotes the frequency of opportunistic executive share reduction. *SCR* denotes supply-chain risk. *SCS* denotes supply-chain resilience. *TE*, *NDE*, *NIE*, and *PM* denote the total effect, natural direct effect, natural indirect effect, and mediation proportion, respectively. *E Value* reports the sensitivity measure for potential unobserved confounding. Columns (1) and (2) examine the pragmatic legitimacy channel. Columns (3) and (4) examine the moral legitimacy channel. Columns (5) and (6) examine the cognitive legitimacy channel. ✓ indicates that the corresponding controls or fixed effects are included. Asterisks report the conventional significance threshold: *** for 1%.

From a mechanistic perspective, the evidence suggests that AI washing is associated with lower external stakeholders' confidence in firms' true value-creation capacity. When firms overstate AI concepts without substantive digital capabilities, outside stakeholders are more likely to regard such technological narratives as exaggerated. Long-term investors place greater weight on sustained growth capacity. They therefore revise downward their assessment of long-term value and strategic execution quality. Business partners place greater weight on operating stability and contract fulfillment. They therefore become more doubtful about operating quality and creditworthiness. As a result, firms become less able

to obtain patient capital and trade credit support. More importantly, patient capital and net trade credit are critical external resources for maintaining organizational resilience. Patient capital supports sustained investment and provides a financial buffer against external shocks. Net trade credit helps ease short-term funding pressure and sustain day-to-day operations. When both decline, the pool of external resources available to the firm contracts. Weaker financial buffers reduce the firm's capacity to absorb shocks. Tighter operating turnover also slows the restoration of normal operations. Accordingly, AI washing reduces organizational resilience by eroding pragmatic legitimacy.

4.4.2. Moral Legitimacy

Moral legitimacy concerns whether corporate conduct is perceived as appropriate and responsible, thereby sustaining the trust foundation of credible governance within the organizational system. This study captures the moral legitimacy mechanism from the perspective of managerial opportunism. Specifically, executives' opportunistic share-reduction amount and the corresponding reduction frequency are selected as mediating variables. The former reflects the extent to which insiders use informational advantages to obtain private benefits. The latter reflects the frequency of opportunistic trading by managers.

As shown in Table 10, AI washing significantly weakens firms' moral legitimacy. The two-step evidence suggests that the estimated coefficient of AI washing on the amount of opportunistic executive share reduction is 20.016 and reaches statistical significance at the 10 percent level. The estimated coefficient on the frequency of share reduction is 2.455 and remains statistically significant at the 5 percent level. As reported in Columns (3) and (4) of Table 11, the CMAverse estimates provide further support. The indirect effect via the amount of opportunistic executive share reduction is -0.005 and reaches statistical significance at the 1 percent level. Its mediation proportion is 1.527%. The indirect effect through the frequency of share reduction equals -0.004 and reaches statistical significance at the 1 percent level. Its mediation proportion is 1.249%. These findings indicate that both the amount and the frequency of opportunistic executive share reduction mediate the impact of AI washing on organizational resilience. In addition, the E values for the two paths are 1.698 and 1.571, respectively. This suggests that the above mediating effects are difficult to attribute entirely to potential unobserved confounders.

From a mechanistic perspective, the results are consistent with the interpretation that AI washing is associated with weaker external stakeholders' recognition of firms' true value-creation capacity. When firms overstate AI capability without substantive digital support, managerial conduct is more likely to be interpreted as opportunistic behavior aimed at catering to the market. AI narratives may generate short-term attention and valuation responses. Opportunistic executive share reduction in this process further amplifies external concerns about governance quality. The larger the amount of share reduction, the stronger the signal of insider rent-seeking. The more frequent the reductions, the stronger the market's doubts about managerial responsibility. As a result, firms' moral legitimacy is undermined. More importantly, organizational resilience is shaped not merely by resource support but also by stable governance and credible decision-making. Managerial opportunism damages investor trust and increases internal governance frictions. Once governance trust declines, firms find it harder to form coordinated responses under external shocks. Accordingly, the results are consistent with the interpretation that AI washing is associated with lower organizational resilience through weaker moral legitimacy, as reflected in greater opportunistic executive share reductions.

4.4.3. Cognitive Legitimacy

Cognitive legitimacy concerns whether a firm's strategic actions can be clearly understood and coherently interpreted by external stakeholders, thereby supporting stable expectations and coordinated relationships across the organizational system. This study captures the cognitive legitimacy mechanism through the lens of supply-chain cognition and coordination. Specifically, supply-chain risk exposure and supply-chain resilience capacity are selected as mediating variables. Supply-chain risk exposure captures the instability of the company's external collaborative network. Supply-chain resilience captures the capacity of the supply-chain system to withstand disruptions and recover operations.

The regression results of Table 10 show that AI washing is associated with weaker observable proxies for cognitive legitimacy. The two-step results indicate that the estimated coefficient of AI washing on supply-chain risk is 1.390 and is significant at the 1 percent level. The estimated coefficient on supply-chain resilience is -17.310 and is also significant at the 1 percent level. As shown in Columns (5) and (6) of Table 11, the CMAverse results provide model-based evidence consistent with a cognitive-legitimacy channel. The indirect effect through supply-chain risk is -0.007 and is significant at the 1 percent level. Its mediation share is 2.129%. The indirect effect through supply-chain resilience is -0.021 and is significant at the 1 percent level. Its mediation share is 6.185%. These findings indicate that supply-chain risk and supply-chain resilience provide evidence consistent with transmission channels linking AI washing to organizational resilience under the maintained mediation assumptions. In addition, the E values are 1.820 and 3.313, respectively. This suggests that these estimates are not easily explained by moderate unobserved confounding.

From a mechanistic perspective, AI washing intensifies external stakeholders' misperception of firms' actual capabilities. When firms portray AI as a key tool for improving supply-chain efficiency without developing the corresponding digital coordination capacity, supply-chain partners find it harder to assess firms' true operating capability. The gap between AI narratives and actual capability increases uncertainty in cooperative relationships. Suppliers and customers may then reassess the firm's delivery capability. As a result, supply-chain risk rises. At the same time, AI narratives unsupported by real capabilities cannot improve supply-chain coordination efficiency. The firm's capability to detect external shocks and take timely responses is also constrained. More importantly, rising supply-chain risk weakens operating stability, while lower supply-chain resilience reduces the firm's ability to restore supply and coordinate cooperation. Once the supply-chain system becomes more fragile, external shocks are more likely to transmit into the firm. Accordingly, the results are consistent with the interpretation that AI washing is associated with lower organizational resilience through weaker cognitive legitimacy, as reflected in higher supply-chain risk and lower supply-chain resilience.

4.5. Heterogeneity Analysis

Given differences in industry characteristics, financing relationships, and management team features, the influence of AI washing on organizational resilience may differ across firms. These factors define important boundary conditions of the organizational system. High-tech industry status affects how strongly external stakeholders interpret AI-related narratives as signals of technological capability and future growth. Bank-firm relationships reflect firms' access to financial resources and the strength of creditor monitoring and financial discipline. Educational heterogeneity in the top management team captures differences in cognitive structure, knowledge background, and internal strategic evaluation. Accordingly, this study examines heterogeneity along three dimensions: high-tech industry status, bank-firm relationships, and top management team educational heterogeneity. This

analysis helps identify how industry, financing, and managerial cognition conditions shape the system-level consequences of narrative-investment misalignment.

4.5.1. Industry Category

Industries differ markedly in the penetration and actual application of AI technologies. This directly shapes how the market interprets firms' AI-related statements. Once industry attributes differ, investors also shift their focus when evaluating firms' AI activities. Market assessment standards, therefore, do not remain the same. Under such conditions, the extent to which AI washing distorts external perceptions may vary across industries, and its effect on organizational resilience may also differ in strength. This issue is especially salient in high-tech industries, where firms' actual AI capability is often harder for outside stakeholders to identify accurately. Misalignment between narrative and capability is therefore less likely to be detected promptly. As a result, the adverse consequences of AI washing are more likely to be amplified and to penetrate more deeply into organizational resilience. Based on this consideration, and following Zhao, et al. [59], the sample is partitioned into high-tech and non-high-tech industry groups based on the industry classification standard in order to further examine industry heterogeneity in how AI washing influences organizational resilience.

Columns (1) and (2) of Table 12 report that the estimated AI-washing coefficients are -0.175 for non-high-tech industries and -0.350 for high-tech industries. In economic terms, a one-standard-deviation increase in *AIW* is associated with a decrease in organizational resilience equivalent to 2.253% of the average annual change in organizational resilience for low-tech firms. For high-tech firms, the corresponding magnitude is 4.176%. The bootstrap coefficient-difference test further confirms that the difference between the two groups is statistically significant, with a *p*-value of 0.000. This result suggests that AI washing may be more damaging in high-tech industries, where AI-related narratives are more closely tied to stakeholders' expectations about technological capability and future growth.

Table 12. Heterogeneity analysis of the effect of AI washing on organizational resilience.

	(1) Low Tech Industry Res	(2) High Tech Industry Res	(3) Low Edu Heterogeneity Res	(4) High Edu Heterogeneity Res	(5) Non-BFR Res	(6) BFR Res
<i>AIW</i>	-0.175^{**} (0.078)	-0.350^{***} (0.086)	-0.254^{***} (0.076)	-0.291^{***} (0.084)	-0.155^{*} (0.090)	-0.356^{***} (0.075)
<i>Constant</i>	-0.000 (0.000)	-0.000 (0.000)	-0.000^{*} (0.000)	-0.000 (0.000)	-0.000^{**} (0.000)	-0.000^{*} (0.000)
Coeff. diff.		0.000		0.000		0.002
Controls	✓	✓	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
<i>N</i>	20,179	15,885	18,554	16,901	14,484	21,280

Note: This table presents subsample evidence on how the AI washing–organizational resilience relationship varies across firm and environmental characteristics. Columns (1)–(2) compare non-high-tech firms with high-tech firms. Columns (3)–(4) split the sample according to whether the top management team exhibits low or high educational heterogeneity. Columns (5)–(6) contrast firms without bank-firm relationships with those maintaining such relationships. All model specifications adjust for the complete set of control variables and absorb firm- and year-level unobserved heterogeneity. ✓ indicates that the corresponding controls or fixed effects are included. Asterisks indicate conventional significance thresholds: *** for 1%, ** for 5%, and * for 10%. Standard errors appear in parentheses.

This finding is closely related to the intrinsic characteristics of high-tech industries. Technological change is rapid in these industries, and outside stakeholders find it more difficult to assess firms' actual AI capability accurately [37]. Once the narrative moves

ahead of the investment, the market is less able to identify this deviation promptly. Misjudgment induced by AI washing is therefore more likely to persist. More importantly, the competitiveness of high-tech firms depends on sustained technological accumulation. Once excessive resources are diverted to concept packaging, the building of real capability is crowded out, and organizational resilience is weakened. If limited resources are allocated excessively to narrative promotion rather than substantive capability building, the adverse effects are more likely to spread quickly to operational adjustment and post-shock recovery [60]. By contrast, technological complexity is lower in non-high-tech industries, so outside stakeholders can usually assess firms' actual operating conditions more directly [61]. The scope for cognitive distortion created by AI washing is therefore more limited. Taken together, AI washing exerts a stronger suppressing effect on organizational resilience in high-tech industries.

4.5.2. Educational Heterogeneity Within the Top Management

Educational heterogeneity within the top management team reflects differences in knowledge background and cognitive structure. Greater educational heterogeneity can broaden the firm's understanding of AI technologies, but it may also increase the difficulty of forming a shared judgment about whether AI-related narratives are supported by substantive investment. When managers differ more substantially in educational backgrounds, their cognitive frames, professional languages, and strategic priorities may diverge, making internal coordination more demanding under conditions of narrative-investment misalignment. By contrast, management teams with lower educational heterogeneity may face fewer internal interpretation and coordination frictions. On this basis, the sample is split into two groups characterized by high and low educational heterogeneity in the top management team to further examine managerial heterogeneity in the relationship between AI washing and organizational resilience.

The estimates reported in columns (3) and (4) of Table 12 indicate that the estimated AI-washing coefficients equal -0.254 and -0.291 for firms in the low- and high-educational-heterogeneity groups, respectively. This indicates that the negative effect of AI washing is stronger among firms with higher educational heterogeneity in the top management team. In economic terms, a one-standard-deviation increase in AIW is associated with a decrease in organizational resilience equivalent to 3.128% of the average annual change in organizational resilience for firms with low educational heterogeneity. For firms with high educational heterogeneity, the corresponding magnitude is 3.616%. The bootstrap coefficient-difference test further confirms that the difference between the two groups is statistically significant, with a p -value < 0.001 . This finding suggests that greater educational heterogeneity may amplify the coordination costs associated with AI washing. When managers have more diverse educational backgrounds, differences in cognitive frames and strategic priorities may make it harder to form a unified assessment of whether AI narratives are supported by substantive investment. As a result, the mismatch between AI narratives and actual AI investment may generate stronger internal coordination frictions and further weaken organizational resilience.

These findings suggest that educational heterogeneity within the top management team may weaken the adverse association between AI washing and organizational resilience. According to the upper echelons theory, executives' educational backgrounds shape their cognitive bases, information-processing patterns, and strategic choices [62]. However, when AI narratives move ahead of substantive AI investment, this diversity may also raise the cost of reaching a common interpretation of the firm's actual AI capability, implementation priorities, and disclosure credibility. Differences in professional training and strategic attention may make it harder for the team to coordinate corrective

responses quickly once the mismatch between narrative and investment emerges. As a result, narrative-investment misalignment is more likely to generate internal coordination frictions and weaken organizational resilience among firms with higher educational heterogeneity in the top management team. Taken together, the results indicate a stronger negative estimated relationship between AI washing and organizational resilience when top management teams have higher educational heterogeneity.

4.5.3. Bank-Firm Relationships

Bank-firm relationships reflect long-term interactions between firms and financial institutions. They provide stable financing support and also strengthen banks' ongoing monitoring of firms' operating conditions. Firms with bank-firm relationships usually rely more heavily on bank credit to maintain working capital and strategic investment. When a gap emerges between firms' AI narratives and actual capabilities, banks are therefore more likely to interpret it as a sign of inadequate disclosure or rising governance risk. Under such conditions, AI washing not only distorts external market perceptions but may also trigger a trust penalty from financial institutions. On this basis, the sample is divided according to whether a firm has a bank-firm relationship to further investigate heterogeneity in the impact of AI washing on organizational resilience across financing relationships.

The negative effect of AI washing is also more pronounced among firms with established bank-firm relationships. In the non-BFR group, the coefficient on *AIW* is -0.155 , whereas in the BFR group, the coefficient is -0.356 . This suggests that firms with stronger bank-firm ties experience a larger decline in organizational resilience when AI narratives exceed substantive AI investment. In economic terms, a one-standard-deviation increase in *AIW* is associated with a decrease in organizational resilience equivalent to 1.889% of the average annual change in organizational resilience for firms without bank-firm relationships. For firms with bank-firm relationships, the corresponding magnitude is 4.461%. The bootstrap coefficient-difference test further confirms that the difference between the two groups is statistically significant, with a p -value of 0.002. This result suggests that bank-firm relationships may amplify the adverse consequences of AI washing. When firms maintain closer financing relationships with banks, AI-related narratives are more likely to influence creditors' expectations and assessments of firm credibility. Once such narratives are not supported by substantive AI investment, the reassessment of credibility may weaken external financial support and reduce organizational resilience.

This result suggests that bank-firm relationships do not necessarily buffer the adverse consequences of AI washing. Instead, firms with such relationships are more vulnerable to the erosion of financial trust. In long-term cooperation, banks accumulate more information about firms' operations and are therefore more likely to detect the capability shortfall and governance risk behind AI washing. Once banks revise downward their assessment of operating quality and creditworthiness, firms may face tighter credit, higher financing costs, or stricter loan review. Because organizational resilience depends on stable funding support, weaker financial backing directly constrains firms' ability to absorb shocks and restore operations. Accordingly, AI washing exerts a stronger suppressing effect on organizational resilience among firms with bank-firm relationships.

4.6. Further Analysis

After documenting a negative estimated association between AI washing and organizational resilience, this study carries out further tests along two dimensions, namely the direction as well as the magnitude of narrative capability deviation. This analysis extends the systems logic by examining whether different forms of alignment between AI narratives and substantive AI investment have asymmetric relationships with resilience. The first case

is a positive deviation, in which AI disclosure exceeds actual AI investment and therefore corresponds to AI washing. The second case is a negative deviation, in which actual AI investment exceeds AI disclosure and reflects a stronger emphasis on substantive capability building. In addition, firms with severe AI washing are identified to test whether a larger narrative capability gap is associated with a larger decline in organizational resilience. This design helps determine whether the baseline findings are driven by directional mismatches between AI narratives and actual investment rather than by AI-related disclosure itself.

Table 13 shows that the estimated coefficient on *AIW_pos* equals -0.347 and reaches statistical significance at the 1 percent level. The coefficient estimated on *AIW_neg* equals 0.469 and is also statistically significant at the 1 percent level. These findings suggest that when AI narratives exceed actual investment, organizational resilience is significantly weakened. By contrast, when actual investment exceeds external narratives, organizational resilience is enhanced. Further, the estimated coefficient on *AIW_high* equals -0.001 and reaches statistical significance at the 1 percent level. This result indicates that, as the gap between AI narratives and actual investment becomes larger, the negative effect on organizational resilience becomes more pronounced.

Table 13. Further analysis of asymmetric AI washing effects and severe AI washing.

	(1) <i>Res</i>	(2) <i>Res</i>	(3) <i>Res</i>
<i>AIW_pos</i>	-0.347^{***} (0.100)		
<i>AIW_neg</i>		0.469^{***} (0.095)	
<i>AIW_high</i>			-0.001^{***} (0.000)
Controls	✓	✓	✓
Firm FE	✓	✓	✓
Year FE	✓	✓	✓
N	35,764	35,764	35,764

Note: This table extends the baseline analysis by distinguishing the direction and intensity of the mismatch between AI narratives and actual AI investment. *Res* denotes organizational resilience. *AIW_pos* captures the positive deviation between AI disclosure and actual AI investment. *AIW_neg* captures the negative deviation between AI disclosure and actual AI investment. *AIW_high* is a dummy variable that equals one for firms with severe AI washing and zero otherwise. Columns (1) and (2) examine the asymmetric effects of positive and negative deviations. Column (3) focuses on the resilience consequences of severe AI washing. ✓ indicates that the corresponding controls or fixed effects are included. Standard errors appear in parentheses. Asterisks indicate conventional significance thresholds: *** for 1%.

The above results further reveal the internal logic through which AI washing affects organizational resilience. The decline in organizational resilience does not arise from AI disclosure itself. It arises from the capability mismatch created when AI narratives exceed actual investment. Firms whose actual investment exceeds external narratives usually possess a stronger technological foundation. Their AI capability is therefore more likely to be translated into operational adjustment and risk response capacity. By contrast, when firms overstate AI concepts without substantive digital capability, external stakeholders are more likely to question the credibility of disclosure and the quality of strategic execution. Severe AI washing further amplifies such doubts. As a result, the firm's legitimacy base is more easily eroded. Its capacity to mobilize resources and recover from shocks also declines.

5. Discussion

This study provides new insight into how AI-related narratives relate to organizational resilience. Prior studies on AI disclosure and AI adoption generally emphasize the informational value of AI-related communication and the productivity implications of

AI investment [5,15]. Our findings are consistent with this literature in showing that AI can support organizational adaptation when it is backed by substantive investment and capability formation. At the same time, they qualify the optimistic view of AI disclosure by showing that AI-related narratives are not necessarily informative signals of real technological capability. The value of digital and general-purpose technologies depends not only on formal adoption but also on complementary intangible investment, workplace organization, and capability reconfiguration [6]. This study extends this argument from AI adoption to AI disclosure by showing that the credibility and investment basis of AI-related narratives matter for organizational resilience. When symbolic AI claims are not supported by substantive AI investment and capability formation, they are associated with weaker coordinated functioning of the organizational system. The key issue is therefore not whether firms talk about AI, but whether AI narratives are supported by credible investment and implementation evidence.

This argument also extends research on institutional decoupling and corporate washing. Institutional theory suggests that organizations may adopt external symbols to satisfy legitimacy expectations while decoupling these symbols from internal practices [27]. Corporate washing research similarly shows that organizations may use favorable external claims even when substantive actions are limited [35,36]. Recent AI washing studies further emphasize the gap between AI-related claims and underlying technological capabilities or investment foundations [3,13]. Our findings confirm the basic logic of this literature: symbolic claims become problematic when they are not supported by substantive action. However, they also extend prior work by showing that AI washing is not only a disclosure or impression-management problem. Because AI technologies are complex, difficult for outsiders to verify, and closely tied to expectations about future competitiveness, AI washing can become a systems-level source of fragility. In this respect, the study moves beyond a narrow capital-market interpretation of washing by showing that AI-related narrative exaggeration is also linked to internal coordination, stakeholder trust, and resilience-related organizational outcomes.

The findings further refine the link between organizational legitimacy and resilience. Legitimacy theory explains why stakeholders support firms whose actions are perceived as desirable, proper, and understandable [27]. Resilience research treats resilience as a capability involving anticipation, coping, recovery, and adaptation, rather than as a single performance outcome [63]. This study connects these two streams more directly. It provides evidence consistent with the view that legitimacy is not only a source of stakeholder approval or market recognition but also an organizational condition that supports the resources, governance credibility, and interorganizational coordination required for resilience. From the perspective of resource dependence theory, resilient firms rely on continued external support, including capital, credit, supply-chain cooperation, and market recognition [28]. From the dynamic capabilities perspective, resilience also requires firms to sense environmental changes, mobilize resources, and coordinate adaptive responses [29]. The mechanism evidence is consistent with the interpretation that AI washing is linked to lower resilience through weaker pragmatic, moral, and cognitive legitimacy. In this way, legitimacy loss is translated into weaker resource buffers, lower governance credibility, and poorer operational coordination.

These patterns refine how the relationship between legitimacy and resilience should be understood. Prior research often treats legitimacy as a source of stakeholder approval and resource access, while signaling theory suggests that the value of corporate disclosure depends on whether signals are credible and supported by costly underlying commitments [27,64]. The evidence in this study connects these ideas by showing that AI-related claims become organizationally consequential when stakeholders question whether they

are supported by real capability formation. This helps explain why the consequences of AI washing are not confined to external perceptions or market reactions. They can enter the organizational system through resource withdrawal, weakened trust in managerial judgment, and lower confidence among interorganizational partners. This interpretation is also consistent with recent stakeholder governance research, which emphasizes that firms create and protect value through credible cooperation with stakeholders and through governance arrangements that sustain joint value creation [65]. The heterogeneity evidence further extends this argument by identifying where credibility is especially important: technological credibility in high-tech industries, financial trust in bank-firm relationships, and shared strategic understanding within cognitively diverse top management teams. In this sense, the findings shift attention from the volume of AI disclosure to the credibility of AI-related claims in supporting coordinated adaptation, a core element of organizational resilience as a capability.

Overall, this study suggests that AI washing should be understood as a systems-level mismatch between technological narratives and capability formation. This view moves beyond treating AI washing as merely exaggerated disclosure or weak communication quality. The main theoretical contribution is to reframe AI disclosure as a credibility-dependent organizational signal rather than a simple indicator of technological orientation. Symbolic AI claims can affect the organizational system by reshaping stakeholder expectations, weakening credibility, and disrupting coordination across financial, governance, and supply-chain relationships. The main insight is that AI disclosure becomes consequential not simply because firms mention AI but because such claims create expectations about capability formation, resource commitment, and adaptive capacity. When these expectations are not supported by substantive investment and implementation evidence, AI narratives may become a source of system-level fragility. At the same time, the findings should be interpreted with appropriate caution. The mechanism variables are observable proxies for legitimacy-related outcomes rather than direct measures of stakeholders' subjective perceptions, and the empirical interpretation relies on the assumptions underlying the AIW measure, DDML, AutoIV, and causal mediation analysis. These issues are addressed in the empirical design and limitations sections.

6. Conclusions, Implications, and Limitations

6.1. Conclusions

This study examines whether and how AI washing is associated with organizational resilience. We conceptualize AI washing as a mismatch between symbolic AI narratives and substantive AI investment and construct a firm-level AI washing index based on the gap between AI disclosure and AI-related investment among Chinese A-share listed firms. The empirical results show that AI washing is negatively associated with organizational resilience under the maintained empirical assumptions. This finding suggests that AI-related disclosure does not necessarily indicate stronger adaptive capacity; when AI narratives are not supported by substantive capability formation, they may instead signal weaker organizational foundations for absorbing shocks, coordinating responses, and adapting under uncertainty.

The mechanism analysis provides evidence consistent with legitimacy-related channels. AI washing is associated with weaker stakeholder support, lower governance credibility, and less stable operational coordination, as reflected in patient capital, trade credit, opportunistic executive share reductions, supply-chain risk, and supply-chain resilience. These variables are interpreted as observable proxies for legitimacy-related outcomes rather than direct measures of organizational legitimacy. Further analyses show that the negative estimated relationship is more pronounced in high-tech industries, firms with

established bank-firm relationships, and firms with higher educational heterogeneity in the top management team.

This study contributes to research on AI disclosure, technology washing, and organizational resilience by shifting attention from the volume of AI-related narratives to their credibility and substantive investment basis. It shows that AI narratives become organizationally consequential when they create expectations about capability formation that are not supported by actual investment or implementation. In this sense, AI washing is not only a disclosure-quality issue but also a systems-level mismatch associated with weaker alignment among technological narratives, resource commitments, stakeholder trust, and coordinated adaptation.

6.2. Managerial Implications

The findings suggest that firms should treat AI disclosure as a capability-backed communication process rather than as a short-term narrative strategy. Managers should ensure that major AI-related claims are supported by verifiable evidence, such as approved budgets, implemented software or hardware systems, technical personnel, data resources, application scenarios, and measurable deployment progress. An internal AI-disclosure review mechanism may help align external communication with actual investment and implementation. Such a mechanism should involve strategy, finance, technology, compliance, and investor-relations teams, rather than leaving AI disclosure solely to communication functions.

For market participants, the results highlight the importance of evaluating the consistency between AI narratives and substantive capability formation. Investors, creditors, and analysts should not treat AI-related disclosure as direct evidence of technological readiness. Instead, they should compare AI narratives with observable indicators such as AI-related software and hardware investment, AI-related hiring, patents or R&D evidence, application scenarios, and operational integration. This consistency check is especially important for high-tech firms and firms with close bank-firm relationships, where AI-related claims may have stronger implications for technological credibility and financial trust.

For regulators and policymakers, the evidence points to the value of improving the transparency and verifiability of corporate AI disclosure. Disclosure guidance could encourage firms to distinguish AI-specific investment from general digitalization expenditure, describe the business scenarios in which AI systems are deployed, report implementation progress, and explain the governance arrangements used to manage AI-related risks. These implications should be interpreted in light of the Chinese listed-firm setting and the identification assumptions of the study, rather than as direct evidence for specific regulatory interventions.

6.3. Limitations and Future Research

This study has three limitations that suggest useful directions for future research. First, the empirical analysis focuses on Chinese listed firms. This setting is well suited to examining AI washing because listed firms provide relatively systematic disclosure and accounting information and because AI-related policy attention and capital-market narratives are especially salient in China. However, the findings may not directly generalize to non-listed firms, firms in other countries, or firms operating under different disclosure and regulatory regimes. Future research could examine whether the organizational consequences of AI washing vary across institutional environments, ownership structures, and disclosure systems.

Second, although the analysis includes baseline controls, firm and year fixed effects, and additional controls for internal control quality and managerial overconfidence, residual

confounding cannot be fully ruled out. Unobserved factors such as strategic foresight, managerial attention, risk culture, organizational slack, or digital transformation governance may influence both AI washing and organizational resilience. Future research could address this issue by using sharper quasi-natural experiments, regulatory or disclosure shocks, or richer organizational data that better separate AI narrative exaggeration from other sources of firm weakness.

Third, the mechanism analysis relies on observable proxy variables for pragmatic, moral, and cognitive legitimacy. Patient capital, net trade credit, opportunistic executive share reductions, supply-chain risk, and supply-chain resilience capture theoretically relevant manifestations of legitimacy, but they do not directly measure stakeholders' subjective legitimacy perceptions. Future research could use surveys, interviews, stakeholder evaluations, or external stakeholder textual data to construct more direct measures of organizational legitimacy and further examine how different stakeholders perceive and respond to AI washing.

Author Contributions: Conceptualization, J.S. and Y.X.; methodology, J.S. and Y.X.; software, J.S. and Y.X.; validation, J.S., Y.X., and J.L.; formal analysis, J.S. and Y.X.; investigation, J.S. and Y.X.; resources, H.S. and Y.X.; data curation, J.S. and Y.X.; writing—original draft preparation, J.S., Y.X., J.L., K.F., H.S. and N.L.; writing—review and editing, J.S. and Y.X.; visualization, J.S. and Y.X.; supervision, Y.X.; project administration, Y.X.; funding acquisition, Y.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (Grant No. 72103082), the Social Science Foundation of Jiangsu Province (Grant No. 25GLB019), the Major Project of Philosophy and Social Sciences Foundation of the Higher Education Institutions of Jiangsu Province (Grant No. 2025SJZD051), the Postgraduate Research & Practice Innovation Program of Jiangsu Normal University (Grant No. 2025XKT1374), the Key Project of National Social Science Foundation of China (Grant No. 23AGL029), and the Humanities and Social Science Fund of the Ministry of Education of China (Grant No. 23YJA910004).

Data Availability Statement: The financial, governance, and firm-level accounting data used in this study were obtained from the CSMAR database, and the annual reports were obtained from CNINFO. Because some source data are subject to third-party database licensing restrictions, the raw commercial data cannot be redistributed. To facilitate replication, the authors will provide the AI-related keyword dictionary, text-processing rules, LLM prompt, classification criteria, representative LLM outputs, manual audit protocol, AI disclosure classification results, AI washing index construction procedure, variable-construction files, and estimation scripts for the DDML, AutoIV, mediation, robustness, and heterogeneity analyses.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial intelligence
AIW	AI washing
Res	Organizational resilience
PC	Patient capital
NTC	Net trade credit
SELLA	Amount of opportunistic executive share reduction
SELLT	Frequency of opportunistic executive share reduction
SCR	Supply-chain risk
SCS	Supply-chain resilience
FPG	Firm pay gap
LLM	Large language model

AutoIV	Automatic instrumental variable
DDML	Double-debiased machine learning
CMA	Causal mediation analysis
TE	Total effect
NDE	Natural direct effect
NIE	Natural indirect effect

Appendix A. Large Language Model-Based AI Disclosure

Appendix A.1. Data and Preprocessing

We collected all the annual reports of the Chinese A-share listed firms from CNINFO. To capture the AI-related contents, we followed Yao, et al. [64] to build an AI-related corpus, including 73 keywords, as shown in Table A1. We then identified all the sentences containing these keywords as the raw data of AI-related information extraction.

Table A1. Keywords of AI disclosure.

Artificial Intelligence	AI Products	AI Chips	Machine Translation
Machine Learning	Computer Vision	Human Computer Interaction	Deep Learning
Neural Networks	Biometric Recognition	Image Recognition	Data Mining
Feature Recognition	Speech Synthesis	Speech Recognition	Knowledge Graph
Smart Banking	Smart Insurance	Human Machine Collaboration	Smart Regulation
Smart Education	Smart Customer Service	Smart Retail	Smart Agriculture
Robo Advisory	Augmented Reality	Virtual Reality	Smart Healthcare
Smart Speakers	Intelligent Voice	Smart Government	Autonomous Driving
Smart Transportation	Big Data Operations	Voiceprint Recognition	Feature Extraction
Driverless Driving	Smart Home	Question Answering System	Facial Recognition
Business Intelligence	Smart Finance	Recurrent Neural Networks	Reinforcement Learning
Intelligent Agent	Smart Elderly Care	Big Data Marketing	Big Data Risk Control
Big Data Analytics	Big Data Processing	Support Vector Machine	Long Short-Term Memory
Robotic Process Automation	Natural Language Processing	Distributed Computing	Knowledge Representation
Intelligent Chips	Wearable Products	Big Data Management	Intelligent Sensors
Pattern Recognition	Edge Computing	Big Data Platforms	Intelligent Computing
Intelligent Search	Internet of Things	Cloud Computing	Augmented Intelligence
Voice Interaction	Smart Environmental Protection	Human Machine Dialogue	Deep Neural Networks
Convolutional Neural Networks	Big Data		

Note: This table reports the AI-related keyword dictionary used to retrieve candidate sentences before LLM-based classification.

Appendix A.2. Task Definition and Label Schema

We define AI-related information as sentence-level disclosures that describe the firm's actual adoption, deployment, development, or construction of AI technologies, systems,

products, or infrastructure in its operations, management, or products and services. To operationalize this definition, we classify each sentence into a binary relevance label, `AI_relevance_label`.

Specifically, `AI_relevance_label` equals 1 if the sentence indicates firm-specific AI use or implementation, such as deploying machine learning or deep learning models; building knowledge graphs or question answering systems; developing or using NLP-based applications such as speech recognition, speech synthesis, machine translation, or human-machine dialogue; adopting AI chips or intelligent computing infrastructure such as edge computing; applying computer vision or biometric identification; implementing industry-specific intelligent systems such as smart banking, smart customer service, or smart risk control; or constructing data-driven intelligent platforms such as big data analytics platforms combined with intelligent computing.

A sentence is labeled 0 if it only provides generic definitions, macro-level narratives, policy background, industry trends, or aspirational statements without an explicit firm-level action, project, system, or implementation. Importantly, keyword presence alone is insufficient. A sentence must express firm-level action or implementation semantics to be considered AI-related.

Finally, the model outputs a self-assessed confidence score in the range from 0 to 1 with two decimal places, reflecting uncertainty in the relevance judgment, together with a brief reason statement capped at 30 characters to facilitate manual auditing.

Appendix A.3. Prompting and Model Inference Setup

We implement sentence-level extraction of AI-related disclosures using the specific large language model, Qwen-max, provided through DashScope's OpenAI-compatible endpoint. Each sentence is submitted as a standalone input, and the model is instructed to behave as an annual report disclosure annotator through a fixed system prompt. The system prompt specifies a binary relevance label indicating whether the sentence describes a firm-level AI application, together with an annotator-style confidence score. To facilitate downstream parsing and auditing, the model is required to return a strict JSON object with three fields: `ai_relevance_label`, `confidence`, and a short reason statement.

Model inference is executed with Qwen-max under deterministic decoding, with temperature set to 0 and `top_p` set to 1. The user message wraps the target sentence and explicitly requests the pre-specified JSON output format. To reduce spurious context inference, the model is instructed to rely only on the current sentence unless it contains explicit anaphora cues such as "the above system" or "this platform." We parse the returned text using a robust JSON loader that extracts the first JSON block if the response contains extra tokens, and we normalize outputs into a fixed schema.

To scale to a large corpus of annual report sentences, we use parallel processing with a thread pool and a global rate limiter to control request frequency. Specifically, we cap the request rate at 8 requests per second using a sliding window limiter and run up to 200 worker threads, which allows high throughput while reducing the likelihood of server-side throttling. Failed requests are retried up to four times with exponential backoff and random jitter. If all retries fail, we return a fallback JSON record that flags the sentence as non-relevant with zero confidence and stores the error message for traceability.

For reproducibility and fault tolerance, we implement checkpointing and atomic writes. Each processed sentence is immediately appended to a checkpoint CSV file containing the row identifier, the input sentence, the parsed label, the confidence score, and the raw model output. The pipeline supports resuming from interruptions by reading the set of completed row identifiers from the checkpoint file and processing only the remaining sentences. In addition, we periodically export an Excel snapshot by writing to a temporary file and then

atomically replacing the output file, which prevents corruption if the process is interrupted during saving.

Appendix A.4. Human Validation

To assess the reliability of the LLM-based sentence-level labels, we draw an audit sample from the full corpus using random sampling. Concretely, we randomly draw 3000 sentences, approximately 1 percent of the full sentence corpus, for subsequent annotation and auditing.

We implement a three-reviewer audit process consisting of two research assistants and one supervising reviewer. As described in Appendix A.3, for each sampled sentence, the LLM has produced a structured output including AI_relevance_label, confidence, and a brief reason. The two research assistants independently examine the original sentence and assign labels to AI_relevance_label following a written coding manual that mirrors the study definition in Appendix A.2. The LLM-based labels are unavailable to the research assistants so that their judgments are not biased.

After the two research assistants complete their independent reviews, we compare their decisions and identify disagreements. To assess inter-annotator reliability, we compute Cohen's kappa between the two independent reviewers. The kappa statistic is 0.771 for the AI relevance task, indicating substantial agreement. These results support the reliability of the human audit benchmark used in the subsequent validation exercise. The supervising reviewer then adjudicates all discrepant cases by re-reading the sentence and applying the coding rules, producing a final consensus label for each sentence. The final audited dataset therefore contains the LLM output and the supervisor-adjudicated ground-truth label, which serves as the benchmark for evaluating LLM labeling quality.

To assess the performance of the LLM-based labeling procedure, we also consider two additional LLMs, namely DeepSeek-V3.2 Thinking and Kimi-K2 Thinking. Qwen-max delivers the strongest performance in the AI relevance task. As reported in Table A2, Qwen-max achieves an accuracy of 0.9557, with a high precision of 0.9640, recall of 0.9480, and F1 score of 0.9540, substantially outperforming Kimi-K2 Thinking and DeepSeek-V3.2 Thinking. Overall, the results indicate that Qwen-max provides the most reliable labels. Considering the much more efficient labeling procedure, we believe that the proposed LLM-based AI-related disclosure extraction is well suited to the task. We do not attempt to make a broader claim about model superiority beyond this application. Instead, the evidence shows that, in our validation sample and under a unified prompting protocol, Qwen-max performs best on the specific task of sentence-level AI disclosure extraction.

To further assess the transparency and reproducibility of the sentence-level classification task, we train a BERT-based supervised benchmark using the 3000 manually audited sentences. We use the supervisor-adjudicated labels as the ground truth and do not use the Qwen-max labels as training labels. Specifically, we follow Wei, et al. [65] and fine-tune bert-base-chinese for the same binary classification task and evaluate its performance on a hold-out test set. As reported in Table A2, the BERT-based benchmark performs reasonably well, indicating that a conventional BERT-based classifier trained on our own human-verified labels can provide a useful validation benchmark. However, all performance metrics of the BERT-based benchmark are lower than those of Qwen-max. This comparison supports our use of Qwen-max as the main classifier while showing that the sentence-level classification task can be independently validated by a transparent supervised benchmark trained on our own human annotations, rather than being driven solely by keyword matching.

To avoid overloading the main validation in Table A2, we report class-specific precision, recall, and F1 scores for Qwen-max in Table A3. These additional results show that

the selected model performs well not only on aggregate metrics but also at the class level. In the AI relevance task, the model achieves near-perfect precision for the positive class, indicating that sentences identified as AI-related are almost never false positives. Although recall is somewhat lower, this pattern suggests that the model sacrifices limited coverage in exchange for a very low false-positive rate, which is desirable in our setting because the main measurement concern is overcounting generic or aspirational statements as firm-specific AI-related disclosure. Table A3 further suggests the disadvantage of the traditional keyword-based approach. Specifically, nearly 60% of the samples are misclassified by the keyword-based approach. For further transparency, we report representative examples of LLM classifications in Table A3 and the full prompt in Appendix A.5.

Accordingly, we measure firms' AI disclosure by the frequency of AI-related sentences in annual reports, defined as the number of sentences labeled as AI-related (AI_relevance_label = 1) divided by the total number of sentences in the annual report.

Table A2. Model performance comparison of the LLM-based labeling procedure.

Model	Accuracy	Precision	Recall	F1 Score
AI_relevance_label				
Qwen-max	0.9557	0.9640	0.9480	0.9540
Kimi-K2 Thinking	0.8290	0.8420	0.8111	0.8185
DeepSeek-V3.2 Thinking	0.8760	0.8745	0.8710	0.8726
BERT classifier	0.8467	0.8506	0.7749	0.8110

Note: The best-performing LLM is highlighted in bold. Accuracy is reported as the overall classification accuracy. BERT refers to the fine-tuned bert-base-chinese classifier. Precision, recall, and F1 score are reported for the AI relevance task.

Table A3. Class-specific validation results for Qwen-max.

Case	Support	Precision	Recall	F1 Score
AI_relevance_label				
Yes	1276	0.9991	0.8966	0.9451
No	1724	0.9288	0.9994	0.9628

Note: Precision, recall, and F1 score are reported at the class level, and support denotes the number of manually annotated sentences in each class.

Having established the reliability of the sentence-level LLM labels, we next assess whether the resulting firm-level measures exhibit the expected construct-level patterns. Specifically, we conduct two additional validation exercises. First, we examine construct consistency by relating our text-based AI measures to substantive AI capacity, an alternative proxy for firms' substantive AI capability. If the disclosure measure captures meaningful firm-level AI-related communication, it should be positively associated with substantive AI capability. By contrast, if AI washing captures divergence between symbolic AI narratives and substantive AI commitment, it should be negatively associated with substantive AI capability.

The results are reported in Table A4, Panel A. Column (1) shows that the firm-level AI disclosure measure is positively and significantly associated with substantive AI capacity. Column (2) shows that AI washing is negatively and significantly associated with substantive AI capacity. This opposite-sign pattern is consistent with the intended interpretation of AI washing. Firms with stronger AI-related disclosure tend to exhibit greater substantive AI capability, whereas firms with a larger narrative–commitment gap tend to exhibit weaker underlying AI capability.

Second, we examine whether AI washing merely captures broader market-level AI concept enthusiasm rather than firm-specific divergence between AI narratives and substantive commitment. To address this concern, Panel B of Table A4 re-estimates the baseline

specification after additionally controlling for region-level AI narrative heat, industry-level AI narrative heat, and both jointly. Across all specifications, the coefficient on AIW remains negative and statistically significant. These results suggest that the negative association between AIW and organizational resilience is not simply driven by broader regional or industry-wide AI narrative heat, but instead reflects firm-specific divergence between symbolic AI narratives and substantive AI commitment. In addition to comparing Qwen-max with alternative LLMs, we further provide a supervised-learning benchmark based on a BERT classifier trained on the manually audited sample. As reported in Appendix A.4, the BERT classifier produces reasonable out-of-sample performance, although it does not outperform Qwen-max.

Table A4. Construct consistency and distinction from market-level AI narrative heat.

Panel A. Consistency tests			
	(1) SAI	(2) SAI	
AI	445,384.226 *** (115,283.687)		
AIW		−802,446.436 *** (76,441.473)	
Constant	3746.716 (5501.526)	−42.332 (155.944)	
N	35,764	35,764	
Panel B. The effect of broader market-level AI narrative heat			
	Res (3)	Res (4)	Res (5)
AIW	−0.343 *** (0.056)	−0.280 *** (0.055)	−0.327 *** (0.056)
Constant	−0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Controls	✓	✓	✓
Firm FE	✓	✓	✓
Year FE	✓	✓	✓
N	35,751	35,764	35,751

Note: Panel A reports construct consistency tests using SAI, an alternative proxy for firms' substantive AI capability. Column (1) regresses SAI on the firm-level AI disclosure measure, and Column (2) regresses SAI on AIW. The negative coefficient on AIW in Column (2) indicates that firms with a larger narrative–commitment gap tend to exhibit weaker substantive AI capability. Panel B examines whether AIW merely reflects broader market-level AI narrative heat. Columns (3), (4), and (5) re-estimate the baseline specification for organizational resilience after additionally controlling for region-level AI narrative heat, industry-level AI narrative heat, and both jointly, respectively. ✓ indicates that the corresponding controls or fixed effects are included. Robust standard errors clustered at the firm level are reported in parentheses. *** denotes statistical significance at the 1% level.

Appendix A.5. Full Prompt

The full task prompt used for model inference is provided below for reproducibility.

You are an annual report disclosure annotator. Your task is to read a single sentence extracted from a listed firm's annual report or related report and determine whether it describes the firm's actual use of artificial intelligence (AI), while also providing your confidence in the judgment.

1. Task Description

For each input sentence, you must produce two outputs:

- Whether the sentence involves the firm's AI application (ai_relevance_label)
- Your subjective confidence in the overall judgment (confidence)

Base your judgment only on the current sentence itself. Do not infer additional context unless the sentence clearly refers to a previous sentence (e.g., “the above system,” “this platform,” etc.).

2. Criteria for Identifying AI Application

In this task, “AI application” is narrowly defined as concrete adoption, deployment, or construction by the firm in its own operations, management, products, or services, including but not limited to:

- Using machine learning, deep learning, knowledge graphs, or related models/systems in credit approval, risk management, anti-fraud, or risk control;
- Using intelligent customer service, intelligent recommendation, natural language processing, speech recognition, or related technologies in customer service, marketing, recommendation, investment advisory, or intelligent Q and A;
- Applying image recognition, facial recognition, biometric identification, or computer vision in scenarios such as risk control, office operations, security, or operational management;
- Deploying or operating the firm’s own AI systems or solutions in smart manufacturing, smart agriculture, smart home, smart healthcare, smart finance, smart banking, and similar scenarios;
- Building or launching specific AI-based platforms, systems, models, engines, chips, or algorithmic modules that serve the firm’s business or management.

The following cases are typically NOT considered AI application disclosure (ai_relevance_label = 0):

- Pure conceptual or definitional statements, such as explanations of “AI,” “deep learning,” “knowledge graphs,” or “industrial internet”;
- Industry level or macro background descriptions, such as “AI is the core driver of the new technological revolution” or “the government strongly supports AI development”;
- Discussions of technology trends or industry blueprints without specifying the firm’s own projects, systems, or applications;
- Mere mentions of AI-related keywords (e.g., “machine learning,” “deep learning,” “smart finance”) used in macro narratives or aspirational slogans without describing the firm’s own practices.

3. Confidence Definition and Range

confidence reflects your subjective certainty about the overall classification;

- Provide a decimal between 0 and 1, rounded to two decimal places;
- Values closer to 1 indicate higher certainty, and values closer to 0 indicate lower certainty;
- Clear and typical sentences may receive 0.85–0.99;
- Vague, incomplete, or highly context-dependent sentences may receive around 0.50 or lower.

4. Output format

Output strictly in the following JSON format. Do not add any extra text or explanation: {"ai_relevance_label": 0 or 1,

“confidence”: floating point number,

“reason”: “a brief explanation of your decision, within 30 characters”

Additional requirements:

- ai_relevance_label must be either 0 or 1
- confidence must be a decimal between 0 and 1 with two decimal places, for example, 0.92
- reason must be short text stating the key basis for the judgment.

We subsequently count the frequencies of sentences labeled as AI-related in each firm's annual report to assess the firm-year AI as follows:

Table A5. Representative examples of LLM classifications.

	Human Label	Qwen-Max Label	Note
OpenAI refers to OpenAI, an artificial intelligence research company established in the United States, whose core mission is to ensure that safe artificial general intelligence (AGI) benefits humanity.	No	No	This sentence only introduces the external company OpenAI and does not involve the focal firm's AI applications.
External environmental influences digital and intelligent transformation. As digital and intelligent technologies represented by artificial intelligence, robotics, the Internet of Things, 5G communications, and big data continue to penetrate all sectors of the economy and society, the world has entered the digital economy era, characterized by ubiquitous connectivity, data-driven development, software-defined systems, platform-based support, and intelligence-led growth.	No	No	This sentence only describes broad technological trends and does not mention the firm's specific AI applications.
In the future, the company will further advance the construction of its big data platform, improve the level of intelligence across R&D, manufacturing, management, and service, and comprehensively enhance its operational capabilities in supply-chain management, logistics, and cost control.	Yes	Yes	This sentence refers to the advancement of a big data platform and intelligent operations, but does not specify concrete projects or outcomes.
In addition, the company has actively expanded into smart home and smart video businesses and has supported clients in developing related solutions.	Yes	Yes	This sentence refers to expansion into AI-related fields such as smart home applications, but does not specify concrete projects or outcomes.
In the Internet of Things segment, the company received the 2017 Automotive Electronics Science and Technology Award for Outstanding Enterprise, the OFweek 2017 IoT Innovative Technology Product Award, the 2017 Global IoT and Smart Services Model Gold Dragon Award, and the 2017 Outstanding IoT Enterprise Award, which strengthened its industry position in IoT communications and the vehicle networking sector. In public security, the company was recognized as one of China's Top Ten Security Brands and one of the Ten Most Influential Comprehensive Security Brands in China in 2017. In smart cities, it received honors such as Recommended Brand for Smart City Development in China and Outstanding Solution for Smart Community Development at the First 2017 Smart Community Awards. In intelligent transportation, it was named among the Top Ten System Integrators by business performance in China's urban intelligent transportation market in 2017 and among the Top Ten Outstanding Enterprises in China's traffic video surveillance industry in 2017.	Yes	No	The LLM classified this sentence as a list of awards without explicit AI technologies or applications. Human coders considered these awards to be related to the firm's AI applications.

Table A5. Cont.

	Human Label	Qwen-Max Label	Note
The number of invention patents granted in fields such as artificial intelligence and the Internet of Things ranks among the highest in the world.	No	Yes	The LLM classified this sentence as relevant because it mentions granted invention patents in AI-related fields and is therefore verifiable. Human coders considered it irrelevant because the statement refers to a national-level macro description rather than the focal firm.

Appendix B. Supply-Chain Risk Perception Measurement and Calculation

The supply-chain risk measure used in this study captures the overall supply chain uncertainty perceived by firms. The construction procedure proceeds as follows. First, annual reports of Shanghai and Shenzhen A-share listed firms are collected and converted into analyzable text. The text is then cleaned, segmented, and standardized. Second, following prior text-based risk measurement studies, this study constructs two dictionaries, namely a supply-chain dictionary and a risk dictionary. Seed words are selected from the literature, and related terms are identified through semantic similarity based on Word2Vec. The candidate terms are further screened through frequency statistics and manual verification. Third, following Wu [48], supply-chain risk is measured by the co-occurrence frequency of supply-chain terms and risk terms within a short text window. The specific formula is as follows:

$$Risk_{it} = \frac{1}{T_{it}} \sum_{w=1}^{T_{it}} \{1[w \in \mathbb{R}] \times 1[|w - r| \leq 15]\}, \quad (A1)$$

where T_{it} denotes the total word count in the annual report of firm i in year t . $w = 1, 2, \dots, T_{it}$ represents the words contained in the annual report. $\mathbb{R}$ denotes the supply-chain dictionary. r denotes the risk word closest to a given supply-chain term. Based on Equation (A1), the frequency with which supply-chain terms and risk terms co-occur within a 15-word context window is used to construct the supply-chain risk perception indicator. To facilitate interpretation, the indicator is multiplied by 100.

Appendix C. Detailed Construction of the Supply-Chain Resilience Index

This appendix reports the detailed construction of the supply-chain resilience index. Following the composite index approach, firm-level supply-chain resilience is measured as a weighted sum of standardized indicators:

$$SCS_{it} = \sum_{k=1}^K \omega_k Z_{kit}, \quad (A2)$$

where SCS_{it} denotes the supply-chain resilience of firm i in year t , Z_{kit} denotes the k -th standardized indicator, and ω_k denotes the corresponding weight. A higher value of SCS_{it} indicates that the firm's supply chain is better able to maintain operations under shocks and to achieve recovery and adjustment more rapidly.

The indicator system includes five first-level dimensions, namely adaptive capacity, resistance capacity, recovery capacity, human capital, and government support. Positive indicators are coded such that higher values imply stronger supply-chain resilience, whereas negative indicators are reverse-coded before aggregation. Table A6 reports the detailed composition of the Z_{kit} indicators.

Table A6. Detailed construction of the Z_{kit} indicators for supply-chain resilience.

Code	First-Level Dimension	Second-Level Indicator	Definition	Sign	Weight
Z_{1it}	Adaptive capacity	Sales level	Operating revenue	+	0.0494
Z_{2it}	Adaptive capacity	Capital management efficiency	Operating revenue/Accounts receivable	+	0.3292
Z_{3it}	Adaptive capacity	Production capacity	Total number of employees on duty	+	0.033
Z_{4it}	Adaptive capacity	Informatization level	Degree of digital transformation	+	0.0208
Z_{5it}	Resistance capacity	Innovation output	Total number of patents independently obtained in the current year	+	0.0635
Z_{6it}	Resistance capacity	Equity multiplier	Total assets/Shareholders' equity	—	0.0001
Z_{7it}	Resistance capacity	Property ratio	Total liabilities/Shareholders' equity	—	0.0001
Z_{8it}	Resistance capacity	Fixed assets	Original value of fixed assets minus accumulated depreciation and impairment provisions	+	0.0507
Z_{9it}	Resistance capacity	Risk-taking level	Three-period rolling standard deviation of industry-adjusted return on assets	—	0.0001
Z_{10it}	Resistance capacity	Supply-chain concentration	Average of the sum of procurement and sales shares associated with the top five suppliers and customers	+	0.0037
Z_{11it}	Recovery capacity	Net profit margin on sales	Net profit/Operating revenue	+	0.0001
Z_{12it}	Recovery capacity	Return on equity	Net profit/Shareholders' equity	+	0.0001
Z_{13it}	Recovery capacity	Inventory turnover	Cost of goods sold/Inventory	+	0.0141
Z_{14it}	Recovery capacity	Current ratio	Current assets/Current liabilities	+	0.0148
Z_{15it}	Recovery capacity	Cash generation capacity	Firm's ability to generate cash flow	+	0.046
Z_{16it}	Recovery capacity	Purchases from suppliers	The number of purchases made by the listed firm from suppliers in the current period	+	0.1731

Table A6. *Cont.*

Code	First-Level Dimension	Second-Level Indicator	Definition	Sign	Weight
Z_{17it}	Recovery capacity	Sales to customers	The amount of sales made by the listed firm to customers in the current period	+	0.1302
Z_{18it}	Human capital	Share of employees with a bachelor's degree or above	Number of employees with a bachelor's degree or above/Total employees	+	0.0105
Z_{19it}	Human capital	Share of R&D personnel	Number of R&D personnel/Total employees	+	0.0164
Z_{20it}	Government support	Government subsidies	Total government subsidies received by the firm	+	0.0444
Z_{21it}	Government support	Income tax expense	Amount of income tax expense	+	0.0001

Note: “+” denotes a positive indicator, for which a higher value indicates stronger supply-chain resilience. “−” denotes a negative indicator, for which a higher value indicates weaker supply-chain resilience; negative indicators are reverse-coded before aggregation.

Appendix D. Additional Details on the AutoIV Analysis

Appendix D.1. Candidate Variables for AutoIV Construction

This section describes the candidate variables used for the AutoIV procedure. Following the logic of data-driven automatic instrumental variable construction, these variables are not directly treated as final instruments. Instead, they are used as candidate inputs from which AutoIV learns instrumental variable representations. The purpose is to extract information that is relevant to AI washing while reducing the direct dependence of the learned representation on organizational resilience.

We construct three groups of candidate variables. The first group captures regional AI narrative pressure. Specifically, we calculate the average AI narrative intensity of other firms located in the same region and year, excluding the focal firm itself. This leave-one-out regional measure reflects the local external narrative environment faced by the firm. When firms operate in regions where AI narratives are more prevalent, they are more likely to face stronger external pressure to highlight AI-related strategies, applications, and technological orientation in their disclosures. Therefore, regional AI narrative pressure is expected to be relevant to firms' AI washing behavior.

The second group captures policy-induced AI narrative incentives. We construct this candidate variable by interacting the pilot zone policy shock with firms' pre-policy AI textual intensity. The policy shock reflects changes in the external institutional environment that increase the salience of AI-related development, while pre-policy AI textual intensity captures firms' prior exposure to AI-related narratives. Firms with stronger pre-policy AI exposure are more likely to respond to policy shocks by intensifying AI-related disclosure. This interaction, therefore, provides candidate variation in AI washing induced by policy attention and pre-existing narrative exposure.

The third group captures the external AI attention shock associated with the emergence of ChatGPT. We construct this candidate variable by interacting a post-ChatGPT indicator with firms' pre-GPT AI textual intensity. The rapid diffusion of ChatGPT substantially increased public, investor, and managerial attention to AI. Firms with stronger AI textual exposure before the ChatGPT shock were more likely to be affected by this external attention shock and to adjust their AI-related narratives accordingly. This variable is therefore

designed to capture externally induced AI narrative pressure arising from the sudden increase in market-wide AI attention.

Taken together, these three groups of candidate variables capture distinct sources of external AI narrative pressure, namely regional narrative diffusion, policy-induced attention, and broad technological attention shocks. They are used as candidate inputs for the AutoIV procedure rather than as final manually specified instruments. The AutoIV framework then learns instrumental variable representations from these candidate variables and uses the learned representations in the deep instrumental variable analysis. This design is consistent with the purpose of AutoIV, which is to reduce reliance on a single manually selected instrument and instead exploit theoretically motivated candidate variables to construct data-driven instrumental representations.

Appendix D.2. Robustness to Alternative Candidate Variable Sets in the AutoIV Framework

To examine whether the AutoIV result is driven by a specific group of candidate variables, we re-estimate the AutoIV model after excluding one candidate variable group at a time. The full candidate set includes regional AI narrative pressure, pilot policy shock exposure, and ChatGPT attention shock exposure. We then construct three alternative candidate sets by excluding regional AI narrative pressure, pilot policy shock exposure, and ChatGPT attention shock exposure, respectively. Across all specifications, the sample, control variables, firm fixed effects, year fixed effects, and AutoIV training settings are held constant.

Table A7 reports the results. The estimated average marginal effects of AIW remain negative across all alternative candidate sets. This pattern indicates that the negative AutoIV estimate is not mechanically driven by any single source of candidate variation. Instead, the regional narrative environment, policy-induced narrative exposure, and ChatGPT-related attention shock provide complementary information for the AutoIV construction. The consistency of the coefficient sign across the leave-one-out candidate group's specifications supports the robustness of the deep instrumental variable evidence.

Table A7. AutoIV robustness checks using alternative candidate-variable sets.

	Excluding Regional AI Narrative	Excluding Pilot Policy Shock	Excluding GPT Shock
AIW AME	−0.498	−0.465	−0.517
First stage R^2	0.998	0.997	0.998
Second stage R^2	0.830	0.832	0.831
Controls	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes

References

1. Malik, F.S.; Terzidis, O. Thriving in turbulence: Resilience and strategic adaptation in global business. *Rev. Manag. Sci.* **2025**, *20*, 2349–2394. [\[CrossRef\]](#)
2. Liang, L.; Li, Y. How does organizational resilience promote firm growth? The mediating role of strategic change and managerial myopia. *J. Bus. Res.* **2024**, *177*, 114636. [\[CrossRef\]](#)
3. Sun, Z.; Wen, Y.; Zhao, L.; Almugren, I.; Galgotia, A. Unveiling AI washing: Bridging corporate technological gaps through a cognitive dissonance lens. *Technol. Forecast. Soc. Change* **2026**, *225*, 124511. [\[CrossRef\]](#)
4. Xing, X.; Zhang, Z.; He, W. AI technology, AI narrative, and firm value. *Technovation* **2026**, *149*, 103349. [\[CrossRef\]](#)
5. Frankel, R.; Jennings, J.; Lee, J. Disclosure Sentiment: Machine Learning vs. Dictionary Methods. *Manag. Sci.* **2021**, *68*, 5514–5532. [\[CrossRef\]](#)
6. Brynjolfsson, E.; Rock, D.; Syverson, C. The Productivity J-Curve: How Intangibles Complement General Purpose Technologies. *Am. Econ. J. Macroecon.* **2021**, *13*, 333–372. [\[CrossRef\]](#)

7. Seele, P.; Schultz, M.D. From Greenwashing to Machinewashing: A Model and Future Directions Derived from Reasoning by Analogy. *J. Bus. Ethics* **2022**, *178*, 1063–1089. [\[CrossRef\]](#)
8. Sajko, M.; Boone, C.; Buyl, T. CEO Greed, Corporate Social Responsibility, and Organizational Resilience to Systemic Shocks. *J. Manag.* **2021**, *47*, 957–992. [\[CrossRef\]](#)
9. Lin, J.Y.; Fu, C. China in the Global Economic Structural Transformation and Upgrading. In *Demystifying the World Economic Development*; Lin, J.Y., Fu, C., Eds.; Springer Nature: Singapore, 2024; pp. 505–593.
10. Liu, H.; Tan, M.; Cao, R. The role of AI in promoting urban sustainable development: Evidence from China's new generation of national AI innovative development pilot zones. *Appl. Geogr.* **2026**, *188*, 103906. [\[CrossRef\]](#)
11. Ma, L.; Qi, Z.; Zhang, X.; Hao, F.; Chi, M. How purpose-specific AI use builds organizational resilience: A dynamic capability perspective. *Technol. Soc.* **2026**, *86*, 103311. [\[CrossRef\]](#)
12. Rana, N.P.; Chatterjee, S.; Dwivedi, Y.K.; Akter, S. Understanding dark side of artificial intelligence (AI) integrated business analytics: Assessing firm's operational inefficiency and competitiveness. *Eur. J. Inf. Syst.* **2022**, *31*, 364–387. [\[CrossRef\]](#)
13. Song, X.; Hou, W.; Ouyang, Z.; Hao, F. AI washing: Strategic disclosure and backlash. *Financ. Res. Lett.* **2026**, *95*, 109684. [\[CrossRef\]](#)
14. Carayannis, E.G.; Dumitrescu, R.; Falkowski, T.; Papamichail, G.; Zota, N.R. Enhancing SME resilience through artificial intelligence and strategic foresight: A framework for sustainable competitiveness. *Technol. Soc.* **2025**, *81*, 102835. [\[CrossRef\]](#)
15. Yang, J. How artificial intelligence improves corporate governance: Evidence from the ERNIE large model. *Financ. Res. Lett.* **2026**, *96*, 109782. [\[CrossRef\]](#)
16. Chen, X.; Liu, W.; Chen, G.; Wang, X.; Zhang, Y. Can artificial intelligence enhance corporate resilience? Empirical evidence from China's A-share listed firms. *Pac.-Basin Financ. J.* **2026**, *97*, 103112. [\[CrossRef\]](#)
17. Hu, Y.; Yan, X.; Duan, C.; Yang, X.; Yang, J. Does Artificial Intelligence Improve the Operational Resilience of Enterprises? Evidence from the AI Innovative Application Pioneer Zone Policy in China. *Systems* **2026**, *14*, 377. [\[CrossRef\]](#)
18. Laufer, W.S. Social Accountability and Corporate Greenwashing. *J. Bus. Ethics* **2003**, *43*, 253–261. [\[CrossRef\]](#)
19. Elhajjar, S.; Itani, O.S. AI washing: A conceptual exploration. *AMS Rev.* **2025**, *15*, 519–538. [\[CrossRef\]](#)
20. Shiller, R.J. *Narrative Economics: How Stories Go Viral and Drive Major Economic Events*; Princeton University Press: Princeton, NJ, USA, 2020.
21. Fotheringham, D.; Wiles, M.A. The effect of implementing chatbot customer service on stock returns: An event study analysis. *J. Acad. Mark. Sci.* **2023**, *51*, 802–822. [\[CrossRef\]](#)
22. Akerlof, G.A. The Market for "Lemons": Quality Uncertainty and the Market Mechanism. In *Uncertainty in Economics*; Diamond, P., Rothschild, M., Eds.; Academic Press: Cambridge, MA, USA, 1978; pp. 235–251.
23. Wang, L.; Xue, K.; Luo, D. Media pressure, corporate governance structure, and disclosure quality. *Int. Rev. Econ. Financ.* **2025**, *97*, 103771. [\[CrossRef\]](#)
24. Jin, S.; Xiong, R.; Peng, H.; Tang, S. ESG performance and private enterprise resilience: Evidence from Chinese financial markets. *Int. Rev. Financ. Anal.* **2025**, *98*, 103884. [\[CrossRef\]](#)
25. Li, J.; Zhang, G.; Wang, X. Supply Chain Concentration and Enterprise Resilience: Evidence from China. *Systems* **2026**, *14*, 386. [\[CrossRef\]](#)
26. Sun, Y.; Liu, G.; Wang, J.; Ma, Y.; Guo, J. Can internal auditing mitigate corporate financialization? *Manag. Audit. J.* **2025**, *40*, 584–610. [\[CrossRef\]](#)
27. Suchman, M.C. Managing Legitimacy: Strategic and Institutional Approaches. *Acad. Manag. Rev.* **1995**, *20*, 571–610. [\[CrossRef\]](#)
28. Pfeffer, J.; Salancik, G. External control of organizations—Resource dependence perspective. In *Organizational Behavior 2*; Routledge: London, UK, 2015; pp. 355–370.
29. Teece, D.J. Business Models, Business Strategy and Innovation. *Long Range Plan.* **2010**, *43*, 172–194. [\[CrossRef\]](#)
30. Chung, C.Y. Do long-term institutional investors enhance corporate long-term value orientation? *Appl. Econ.* **2025**, *57*, 8839–8855. [\[CrossRef\]](#)
31. Wang, Z.; Zhao, L.; Shao, Y.; Wen, X. Reputation compensation for incentive alignment in a supply chain with trade credit under information asymmetry. *Ann. Oper. Res.* **2023**, *331*, 581–604. [\[CrossRef\]](#)
32. He, G. How do insider trading incentives shape nonfinancial disclosures? Evidence from product and business expansion disclosures. *Rev. Quant. Financ. Account.* **2023**, *60*, 147–194. [\[CrossRef\]](#)
33. Wang, S.; Zhang, H. Enhancing environmental, social, and governance performance through artificial intelligence supply chains in the energy industry: Roles of innovation, collaboration, and proactive sustainability strategy. *Renew. Energy* **2025**, *245*, 122855. [\[CrossRef\]](#)
34. van der Vegt, G.S.; Essens, P.; Wahlström, M.; George, G. Managing Risk and Resilience. *Acad. Manag. J.* **2015**, *58*, 971–980. [\[CrossRef\]](#)
35. Marquis, C.; Toffel, M.W.; Zhou, Y. Scrutiny, Norms, and Selective Disclosure: A Global Study of Greenwashing. *Organ. Sci.* **2016**, *27*, 483–504. [\[CrossRef\]](#)

36. Walker, K.; Wan, F. The Harm of Symbolic Actions and Green-Washing: Corporate Actions and Communications on Environmental Performance and Their Financial Implications. *J. Bus. Ethics* **2012**, *109*, 227–242. [[CrossRef](#)]
37. Zhang, J.; Huang, W.; Weng, Z.; Ma, J.; Wang, A. Words and deeds in AI transformation: CEO IT background and the rhetoric–substance gap. *Technol. Soc.* **2026**, *86*, 103336. [[CrossRef](#)]
38. Gilardi, F.; Alizadeh, M.; Kubli, M. ChatGPT outperforms crowd workers for text-annotation tasks. *Proc. Natl. Acad. Sci. USA* **2023**, *120*, e2305016120. [[CrossRef](#)] [[PubMed](#)]
39. Williams, T.A.; Gruber, D.A.; Sutcliffe, K.M.; Shepherd, D.A.; Zhao, E.Y. Organizational Response to Adversity: Fusing Crisis Management and Resilience Research Streams. *Acad. Manag. Ann.* **2017**, *11*, 733–769. [[CrossRef](#)]
40. Hillmann, J.; Guenther, E. Organizational Resilience: A Valuable Construct for Management Research? *Int. J. Manag. Rev.* **2021**, *23*, 7–44. [[CrossRef](#)]
41. Ortiz-de-Mandojana, N.; Bansal, P. The long-term benefits of organizational resilience through sustainable business practices. *Strateg. Manag. J.* **2016**, *37*, 1615–1631. [[CrossRef](#)]
42. Martins, A.M.; Cró, S. Airline stock markets reaction to the COVID-19 outbreak and vaccines: An event study. *J. Air Transp. Manag.* **2022**, *105*, 102281. [[CrossRef](#)] [[PubMed](#)]
43. Wiatt, R.D.; Lee, Y.G.; Marshall, M.I.; Zuiker, V.S. The Effect of Cash Flow Problems and Resource Intermingling on Small Business Recovery and Resilience After a Natural Disaster. *J. Fam. Econ. Issues* **2021**, *42*, 203–214. [[CrossRef](#)] [[PubMed](#)]
44. Hu, X.; Lin, D.; Tosun, O.K. The effect of board independence on firm performance—New evidence from product market conditions. *Eur. J. Financ.* **2023**, *29*, 363–392. [[CrossRef](#)]
45. Neukirchen, D.; Engelhardt, N.; Krause, M.; Posch, P.N. Firm efficiency and stock returns during the COVID-19 crisis. *Financ. Res. Lett.* **2022**, *44*, 102037. [[CrossRef](#)] [[PubMed](#)]
46. Dieperink, H.; Adriaanse, J.; Dechesne, M. Predicting viability of small businesses on the edge of failure. *J. Small Bus. Manag.* **2025**, *63*, 2422–2454. [[CrossRef](#)]
47. Bebchuk, L.; Cohen, A.; Ferrell, A. What Matters in Corporate Governance? *Rev. Financ. Stud.* **2009**, *22*, 783–827. [[CrossRef](#)]
48. Wu, D. Text-Based Measure of Supply Chain Risk Exposure. *Manag. Sci.* **2023**, *70*, 4781–4801. [[CrossRef](#)]
49. Bruckler, M.; Wietschel, L.; Messmann, L.; Thorenz, A.; Tuma, A. Review of metrics to assess resilience capacities and actions for supply chain resilience. *Comput. Ind. Eng.* **2024**, *192*, 110176. [[CrossRef](#)]
50. Chernozhukov, V.; Chetverikov, D.; Demirer, M.; Duflo, E.; Hansen, C.; Newey, W.; Robins, J. Double/debiased machine learning for treatment and structural parameters. *Econom. J.* **2018**, *21*, C1–C68. [[CrossRef](#)]
51. Yuan, J.; Wu, A.; Kuang, K.; Li, B.; Wu, R.; Wu, F.; Lin, L. Auto IV: Counterfactual Prediction via Automatic Instrumental Variable Decomposition. *ACM Trans. Knowl. Discov. Data* **2022**, *16*, 74. [[CrossRef](#)]
52. Xu, Y.; Xu, L.; Shen, Y.; Fan, Z. Exploring the effect of digital transformation on firm resilience: Evidence from China. *J. Asian Econ.* **2024**, *95*, 101812. [[CrossRef](#)]
53. Shi, H.; Xia, Y.; Zong, Z.; Hua, Y.; Sun, J.; Chen, X. Does Data Asset Information Disclosure Mitigate Supply Chain Risk? Causal Evidence from Double-Debiased Machine Learning. *Systems* **2025**, *13*, 844. [[CrossRef](#)]
54. Ahrens, A.; Hansen, C.B.; Schaffer, M.E.; Wiemann, T. ddml: Double/debiased machine learning in Stata. *Stata J. Promot. Commun. Stat. Stata* **2024**, *24*, 3–45. [[CrossRef](#)]
55. Tang, H.; Tong, M.; Chen, Y. Green investor behavior and corporate green innovation: Evidence from Chinese listed companies. *J. Environ. Manag.* **2024**, *366*, 121691. [[CrossRef](#)] [[PubMed](#)]
56. Chen, H.; Yang, D.; Zhang, X.; Zhou, N. The Moderating Role of Internal Control in Tax Avoidance: Evidence from a COSO-Based Internal Control Index in China. *J. Am. Tax. Assoc.* **2020**, *42*, 23–55. [[CrossRef](#)]
57. Twardawski, T.; Kind, A. Board overconfidence in mergers and acquisitions. *J. Bus. Res.* **2023**, *165*, 114026. [[CrossRef](#)]
58. Gupta, A. R&D and firm resilience during bad times. *Int. J. Ind. Organ.* **2026**, *105*, 103246. [[CrossRef](#)]
59. Zhou, W.; Zhang, Y.; Li, X. Artificial intelligence, green technological progress, energy conservation, and carbon emission reduction in China: An examination based on dynamic spatial Durbin modeling. *J. Clean. Prod.* **2024**, *446*, 141142. [[CrossRef](#)]
60. Hambrick, D.C.; Mason, P.A. Upper Echelons: The Organization as a Reflection of Its Top Managers. *Acad. Manag. Rev.* **1984**, *9*, 193–206. [[CrossRef](#)] [[PubMed](#)]
61. Duchek, S. Organizational resilience: A capability-based conceptualization. *Bus. Res.* **2020**, *13*, 215–246. [[CrossRef](#)]
62. Connelly, B.L.; Certo, S.T.; Ireland, R.D.; Reutzel, C.R. Signaling Theory: A Review and Assessment. *J. Manag.* **2011**, *37*, 39–67. [[CrossRef](#)]
63. Bridoux, F.; Stoelhorst, J.W. Stakeholder Governance: Solving the Collective Action Problems in Joint Value Creation. *Acad. Manag. Rev.* **2020**, *47*, 214–236. [[CrossRef](#)]

64. Yao, J.; Zhang, K.; Guo, L.; Feng, X. How does artificial intelligence improve firm productivity? Based on the perspective of labor skill structure adjustment. *J. Manag. World* **2024**, *40*, 101–116.
65. Wei, X.; Zhang, Z.; Zhang, M.; Chen, W.; Zeng, D.D. Combining Crowd and Machine Intelligence to Detect False News on Social Media. *Manag. Inf. Syst. Q.* **2022**, *46*, 977–1008. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.