AI Sovereignty

Zhenyu Gao, Wenxi Jiang, Haohan Ren, Kemin Wang, Yuezhi Wu*

November 2025

Abstract

We examine how state power shapes corporate technology adoption. Using novel, high-frequency data from investor-firm dialogues and the staggered releases of ChatGPT (foreign) and DeepSeek (domestic) as shocks, we build monthly, model-specific adoption measures in China. We find sovereignty-exposed firms delay foreign adoption after ChatGPT but accelerate domestic adoption after DeepSeek. This behavior, driven by data security concerns, is stronger under high state influence and weaker in competitive markets. Markets reward firms for this strategic alignment. We provide first systematic evidence that geopolitical forces are creating divergent adoption paths for a frontier technology, with state-influenced firms acting as key agents.

* Zhenyu Gao and Wenxi Jiang are at the Department of Finance, CUHK Business School, The Chinese University of Hong Kong, Shatin, Hong Kong; and Haohan Ren, Kemin Wang, and Yuezhi Wu are at the Department of Finance, School of Management, Fudan University, Shanghai, China. Authors' contact information: Gao: gaozhenyu@baf.cuhk.edu.hk; Jiang: wenxijiang@baf.cuhk.edu.hk; Ren: haohanren@fudan.edu.cn; Wang: wangkm@fudan.edu.cn; and Wu: wuyz23@m.fudan.edu.cn.

1. Introduction

The diffusion of transformative technologies reflects not only economic efficiency and technical merit but also political forces. This dynamic is evident in the global race to adopt generative AI (GenAI), unfolding amid intensifying U.S.–China rivalry. Within this context, AI sovereignty—the imperative for nations to secure its own AI ecosystems—has emerged as a strategic priority.[1]

We examine how this imperative manifests at the corporate level, asking whether firms aligned with state interests systematically favor domestic technologies. In doing so, we respond to Acemoglu's (2025) call in his Nobel lecture to investigate how political incentives shape technology choices and innovation trajectories. Our analysis provides the first systematic evidence of AI sovereignty in practice, showing how geopolitical forces drive divergent paths in corporate adoption of frontier technologies.

China offers a compelling setting for our study, where state interests and market forces are deeply intertwined. The staggered global releases of two landmark large language models provide a powerful quasi-natural experiment. The launch of OpenAI's ChatGPT API in early 2023 created a shock to the availability of leading foreign GenAI[2], while the subsequent release of DeepSeek in early 2025 introduced a competitive, domestically developed alternative. This setting allows us to ask how different firms respond to these shocks and, in particular, whether firms with greater "sovereignty exposure" (i.e. those more closely tied to the Chinese state's interests or less dependent on U.S. technology) exhibit a systematically different adoption pattern compared to other firms.

Empirically testing these ideas requires a timely and detailed view of technology adoption at the firm level. We construct a novel dataset of firm-level GenAI adoption using textual dialogues from the *Investor Interaction Platforms (IIPs)* maintained by China's stock exchanges. These IIPs are online forums where investors publicly pose questions to listed companies and companies are obligated to respond, typically within

[1] Nvidia (2024) defines the sovereign AI as a nation's capabilities to produce artificial intelligence using its own infrastructure, data, workforce and business networks. Mügge (2024) defines AI Sovereignty as "the need for control, as necessary to develop and deploy AI technologies, of computing capacity and data storage, of access to human resources and potentially proprietary knowledge to build AI applications, and of training data". The benefits of AI sovereignty include the fact that "it codifies your culture, your society's intelligence, your common sense, your history – you own your own data" (Huang, 2024), while also enhancing competitiveness (Mügge, 2024). Leaders around the world—not only in the U.S. and China, but also across Europe, the Middle East, and Asia—are showing interest in developing AI sovereignty (Nvidia, 2025).

[2] While OpenAI has restricted direct GPT access within China, Microsoft Azure has been delivering authorized and reliable GPT services to enterprise clients. See e.g., https://azure.microsoft.com/en-us/blog/general-availability-of-azure-openai-service-expands-access-to-large-advanced-ai-models-with-added-enterprise-benefits/

days. Crucially for our purposes, investors have frequently inquired whether firms are using or planning to use specific AI models such as ChatGPT or DeepSeek, prompting firms to confirm or deny usage and sometimes to elaborate on integration plans. We collect over 28,000 question-and-answer dialogues from January 2022 to June 2025 that discuss GenAI technologies, covering 3,656 unique firms (about 70% of all listed companies and nearly 90% of total market capitalization in China). We leverage these detailed dialogues to identify if and when each firm adopts GenAI, and importantly, which exact AI models are involved.

To systematically extract adoption information from the unstructured text, we employ a state-of-the-art large language model (LLM) to classify each firm's response. The LLM is instructed to determine whether the company has deployed GenAI (either in practice or in planned initiatives) and to identify which specific models are being adopted or not adopted. This approach yields a monthly panel of adoption indicators for each firm, including a general GenAI adoption flag and model-specific flags (e.g., indicators for adopting GPT, DeepSeek, Claude, etc.). We treat adoption as persistent once it occurs, as new generative AI releases typically trigger immediate yet persistent adoption spikes (Fradkin, 2025). We conduct several validation checks confirming that our text-derived measure aligns with real actions: firms flagged as adopters in the Q&A data are more likely to mention AI in official filings, register AI algorithms with regulators, hire for AI-related roles, and they experience positive stock price reactions upon disclosing adoption. These patterns bolster confidence that the measure reflects genuine technology deployment rather than mere rhetoric.

Armed with this granular dataset, we compare changes in adoption rates before and after each event between firms with high versus low exposure to state sovereignty concerns. We proxy a firm's "sovereignty exposure" using several measures: whether the firm is state-owned, whether it operates in a state-designated strategic industry (e.g., energy, telecommunications, defense), the degree to which the firm's technologies are decoupled from U.S. intellectual property (using a patent-based index of reliance on U.S. patents (Han et al. 2024), and the firm's stock return sensitivity to U.S.–China geopolitical tensions (Rogers et al. 2024). We combine these related proxies into a composite Sovereignty Index to capture the broad concept of alignment with national interests. Intuitively, firms scoring high on this index include state-owned or policy-

oriented enterprises and those less tied into U.S.-dependent technologies. These are the firms we expect to be at the forefront of China's AI sovereignty campaign.

Our analysis yields four main findings. First, we document divergent adoption paths for high- versus low-sovereignty firms, consistent with a strong AI sovereignty effect. Specifically, we show that firms with high "sovereignty exposure" were significantly less likely to adopt GenAI following the release of ChatGPT. This reluctance, however, reversed dramatically with the advent of a domestic alternative. These same sovereignty-exposed firms were significantly *more* likely to adopt GenAI after the launch of DeepSeek, effectively "catching up" by embracing a domestic champion. This pattern is robust to a battery of tests, including the use of model-specific adoption outcomes, which sharpens identification by directly linking the adoption of a specific model to its release event.

We also test plausible alternative explanations for these adoption patterns. The observed preference for DeepSeek could stem from its open-source license, which offers greater customizability and data control compared to the closed-source ChatGPT, rather than its domestic origin. However, we find that sovereignty-exposed firms exhibit a similar reluctance to adopt other foreign models regardless of their licensing, including both the open-source LLaMA and the closed-source Gemini.

Another alternative is a "low-cost hypothesis," suggesting that firms prefer DeepSeek for its lower price. We test this by controlling for various proxies of firm financial constraints and cost sensitivity. Our main results remain robust, indicating that cost considerations are not the primary driver of the observed behavior. By ruling out these alternatives, we strengthen the evidence that the divergent adoption paths are a direct manifestation of AI sovereignty.

Second, we find that the magnitude of this sovereignty-driven gap is shaped by a state–market trade-off. On one hand, the preference for domestic AI is most pronounced among firms operating in environments where state influence is strongest. For example, the adoption divergence is amplified for firms headquartered in regions with heavier government involvement in the economy (such as Beijing) and for those whose top executives have government backgrounds. In provinces with above-median government spending as a share of GDP – a proxy for the government's role in the local economy – sovereignty-leaning firms show almost no uptake of ChatGPT but a large surge in DeepSeek adoption, relative to other firms. On the other hand, market forces

can temper the role of sovereignty. In highly competitive industries (e.g., those with low concentration ratios), even politically connected firms face greater pressure to use the best available technology (often foreign) lest they fall behind their rivals. We observe much smaller differences in adoption behavior between high- and low-sovereignty firms in sectors characterized by intense competition. In less competitive, more oligopolistic sectors, however, the divergence due to state influence is stark. These interactions suggest that while political imperatives significantly sway firm behavior, their impact is modulated by market discipline: firms in cutthroat markets cannot afford to ignore efficiency advantages, whereas firms with sheltered market positions have more latitude to heed state directives.

Third, we explore the mechanisms behind the defense of AI sovereignty and find that data security is a primary driver. Generative AI systems often require sending data to external servers or models, raising the risk of sensitive information leakage. Chinese firms handling critical data – for instance, those in finance, healthcare, or defense – face strict domestic regulations on data export and cyber security. We find that among firms with high sovereignty exposure, those that manage more data assets or explicitly acknowledge higher data breach liabilities in their reports are the least likely to engage with foreign AI and the most inclined to wait for domestic solutions. The political priority of safeguarding data sovereignty makes firms wary of foreign AI models that could inadvertently funnel proprietary or sensitive data out of China. Furthermore, even when high-exposure firms do adopt a foreign model like ChatGPT, our textual analysis of their disclosures shows they tend to confine its use to low-risk domains (e.g. customer service FAQ) rather than core R&D or proprietary operations – a nuance indicating risk mitigation behavior. In contrast, such firms freely deploy domestic AI in a wider range of applications, reflecting greater trust in locally governed AI.

Finally, we examine how investors react to these strategic choices. An event study of stock returns around adoption announcements reveals that the market prices AI sovereignty. While GenAI adoption is, on average, viewed positively, the reaction is highly conditional on the firm's sovereignty exposure and the origin of the technology. The market responds significantly more favorably when sovereignty-exposed firms announce the adoption of a domestic model compared to a foreign one. This suggests that investors recognize the long-term strategic value of aligning with national priorities

in an era of escalating technological rivalry, rewarding firms that contribute to building a self-reliant domestic AI ecosystem.

These findings contribute along several dimensions. First, it advances the political economy of technology by showing that state incentives shape not only the level but the direction and timing of corporate AI adoption, in line with calls for research on institutional feedback to technology choices. Second, it complements work on protectionism and geoeconomic tensions by documenting a non-tariff, firm-level channel—technology selection—through which geopolitical rivalry reconfigures production possibilities. Third, it adds to research on AI diffusion by introducing non-market drivers—government ownership, regulated-industry status, geopolitical detachment—as key determinants of model-specific adoption, beyond task composition, skills, and productivity. Fourth, methodologically, it demonstrates the use of the IIP dialogues to recover high-frequency, model-level adoption measures validated against revealed corporate actions and priced by financial markets.

Literature Review Our paper contributes to a growing literature at the intersection of artificial intelligence (AI) and political economy. Acemoglu (2025) highlights that political incentives shape technology adoption and calls for more research on the feedback loop between institutions and technological choices. In the Chinese context, government data resources and regulatory power have been shown to accelerate innovation, which in turn reinforces governance capabilities (Beraja, Peng, Yang, and Yuchtman, 2024; Beraja, Kao, Yang, and Yuchtman, 2023; Beraja, Yang, and Yuchtman, 2023). Recent theoretical work also emphasizes the role of data flows and localization policies in shaping global technology competition (Chang, Cong, Wang, and Zhang, 2025; Farboodi and Veldkamp, 2021). Our study provides insights into how a novel political factor in the era of generative AI—namely, the defense of AI sovereignty—shapes the diffusion of these technologies.

Our study also contributes to the literature examining the economic impacts of geopolitical conflicts (Clayton, Coppola, Maggiori and Schreger, 2025; Clayton, Maggiori and Schreger, 2025; Clayton and Maggiori, 2025). Han, Jiang, and Mei (2024) examine the evolving dynamics of U.S.–China technological decoupling and document how policy interventions reshape bilateral dependencies and innovation trajectories. Cen, Fos, and Jiang (2024) show that foreign industrial policies create significant

negative spillovers for U.S. firms. Crosignani, Han, Macchiavelli and Silva (2024) find that export controls aimed at safeguarding technological leadership by the U.S. hurt performance of domestic firms. Canayaz, Erel, Gurun, and Wu (2025) find that restricting high-skilled immigration reduced efficiency and constrained the domestic semiconductor industry. To the extent that the defense of AI sovereignty represents a form of geopolitical conflict, our study provides novel evidence on the impacts of such conflicts on technological development.

A related strand of work examines the strategic behaviors of state-owned or politically-connected enterprises in international conflicts. Benguria and Saffie (2021) find that SOEs strategically adjusted import patterns during the U.S.–China trade war. Chen, Hsieh, and Song (2024) document that the implementation of non-tariff barriers varies with the state ownership. Relatedly, Cen, Cohen, Wu, and Zhang (2025) show that politically-connected firms expanded trade during geopolitical shocks such as the U.S.–China trade war and the Russia–Ukraine conflict. We extend this line of research by showing that the strategic role of SOEs goes beyond trade and sanctions, shaping their decisions in adopting cutting-edge technologies such as generative AI.

Our study also relates to research on the determinants and consequences of AI adoption. Bick et al. (2024) and Bonney et al. (2024) document the rapid diffusion of GenAI across individuals and firms, shaped by education, task composition, and industry characteristics. Fradkin (2025) shows that new GenAI releases drive immediate but persistent adoption spikes, while Schubert (2025) highlights complementarities between remote work intensity and AI usage. In terms of consequences, Babina et al. (2024) find that AI-related hiring predicts productivity and innovation gains, and Eisfeldt, Schubert, Taska, and Zhang (2023) document that greater exposure to generative AI drives significant firm value appreciation. Relative to this literature, we provide novel evidence on model-specific adoption patterns and highlight the role of non-economic motives—such as alignment with state priorities—in shaping firms' adoption decisions, answering to the call over more research at the intersect of Transformative AI (TAI) and geoeconomics (Brynjolfsson, Korinek, and Agrawal, 2025).

Finally, we contribute to the methodological literature applying large language models (LLMs) in finance. Studies such as Jha, Qian, and Weber (2024) and Sheng, Sun, and Zhang (2024) show that LLMs can extract economic insights from

unstructured text, improving forecasting and decision-making in corporate and asset management contexts. In this paper, we apply LLMs to the uniquely granular disclosure environment of China's Investor Interaction Platforms (IIPs), enabling high-frequency identification of both the timing and specificity of GenAI adoption at the firm level.

The remainder of the paper is organized as follows. Section 2 describes the data sources, the construction of the GenAI adoption measure, and summary statistics. Section 3 validates our GenAI adoption measure. Section 4 presents the empirical results on differential adoption patterns and underlying mechanisms, as well as the market reaction analysis. Section 5 concludes with a discussion of the implications of our findings.

2. Data and Measures

2.1 Measure of Firm-level AI Adoption

2.1.1 Investor Interaction Platforms

We construct a firm-level measure of generative AI (GenAI) adoption based on dialogues between investors and listed companies on China's Investor Interaction Platforms (IIPs). These platforms, maintained by the Shanghai Stock Exchange (SHSE) and Shenzhen Stock Exchange (SZSE), serve as official channels for communication between investors and publicly listed firms in China. The SZSE platform Hu Dong Yi (in Chinese, 互动易) was launched in 2010, and the SHSE platform e Hu Dong (in Chinese, e 互动) in 2013. On both platforms, investors submit questions that undergo a brief review before being published. Company participation is de facto mandatory: exchanges require firms to respond to all investor questions, monitor compliance, and may issue warnings or downgrade investor relations (IR) rankings for non-responsiveness, while rewarding firms with high responsiveness. Prior research has demonstrated that these IIPs serve as an important source of firm-level information for investors, and that the dialogues on the platforms offer valuable insights into both investor behavior and corporate disclosure practices (Lee and Zhong, 2022; Wong et al., 2024).

We construct our GenAI adoption measures using dialogues from the IIPs rather than traditional disclosure sources such as annual reports or earnings calls for two main reasons. First, annual reports are released only once a year and primarily provide retrospective, aggregated accounts of technological investments. Earnings calls,

although somewhat more frequent, focus mainly on financial metrics and offer limited detail on technological initiatives. In contrast, IIP dialogues occur in real time, as investors pose questions and firms respond within days. This high-frequency setting enables us to capture monthly adoption dynamics from January 2022 to June 2025, reflecting the rapid evolution of GenAI technologies (e.g., the swift uptake of models like DeepSeek shortly after launch).

Second, IIPs facilitate investor-initiated inquiries that elicit targeted and detailed firm responses about GenAI adoption—information that is often absent or vaguely discussed in conventional disclosures. Investors frequently raise explicit questions concerning the deployment of specific AI models (e.g., "Has your firm integrated ChatGPT or DeepSeek into its operations?"), which compels firms to disclose concrete information about their engagement with GenAI. This enables the construction of model-specific adoption measures, which is typically infeasible when relying on financial reports or job postings, where AI-related information tends to be generic, sparse, and inconsistently reported across firms. Table A6 provides three representative examples of such inquiries and the corresponding firm responses.

2.1.2 Construction of AI Adoption Measure

Our primary objective is to construct a reliable firm-level measure of generative AI (GenAI) adoption. This process involves two main stages: (1) collecting GenAI-related firm disclosures from the IIPs, and (2) employing a large language model to analyze and classify these disclosures in a structured manner.

We start with the complete set of investor-firm dialogues posted on the IIPs from January 2022 to August 2025. To identify discussions specifically related to modern generative AI, we filter this extensive dataset for dialogues where either the investor's question or the firm's response contains at least one keyword related to GenAI. We construct the keyword list of leading generative AI models featured on prominent LLM leaderboards (e.g., Hugging Face Open LLM Leaderboard and LMSYS Chatbot Arena) along with state-of-art domestic ones introduced after 2022.[3] We believe this list

[3] Specifically, the keywords used are: OpenAI, GPT, Sora, LLaMA, Claude, Gemini, DeepMind, Grok, DALL·E, Midjourney, Stable Diffusion, T5, Pika, CLIP, Bloom, and PaLM. The Chinese models include DeepSeek (and its Chinese name "深度求索"), 文心一言 (ERNIE), Ernie, Kimi, 月之暗面 (Moonshot), 豆包 (Doubao), 千问 (Qwen), ChatGLM, 盘古 (Pangu), 天工 (Tiangong), 星火 (Spark), 智谱 (Zhipu), 九天 (Jiutian), 百川 (Baichuan), 天问 (Tianwen), 昇腾 (Ascend), 西湖心辰 (Westlake Xinchen), 紫东太初 (Zidong Taichu), and 可灵 (Keling). To ensure our analysis focuses on GenAI and avoids capturing earlier forms of artificial intelligence,

captures the majority of GenAI models launched in the Chinese and U.S. markets. In total, we identify 28,822 unique GenAI-related dialogues from 3,656 distinct firms. As shown in Figure 1, the GenAI questions rises after major GenAI-related events such as the launches of ChatGPT and DeepSeek services. The timely rise in dialogues confirms that the data source can provide valid GenAI adoption measures at high frequency.

[Insert Figure 1 here]

After collecting the relevant dialogues, we employ DeepSeek V3, a state-of-the-art large language model developed by DeepSeek, to analyze the content of these filtered dialogues. DeepSeek V3 is particularly suited for this task due to its strong performance in natural language understanding, especially in Chinese contexts, and its ability to handle nuanced financial and technological discussions.

For each dialogue, we input the firm's response (reply_content) into a carefully designed prompt that guides the model to classify the firm's GenAI adoption status and extract specific model mentions. The prompt is structured to incorporate financial literacy and AI expertise, ensuring consistent and reliable outputs. It defines "adoption" broadly to include planned initiatives (e.g., Bonney et al., 2024), fine-tuning activities, integration into business operations, or full deployment. The full prompt used is provided in the Appendix (translated from Chinese).

From these classifications, we derive two firm-level GenAI adoption measures. The first one is a binary variable, $AI_Adoption_{i,t}$, which equals 1 if the firm's response in any dialogue during or prior to month t indicates adoption, and 0 otherwise. Adoption is treated as persistent once identified, reflecting the cumulative nature of technology integration. Note that Fradkin (2025) documents that new GenAI releases tend to trigger persistent adoption spikes, providing justification for our variable definition. The second is model-specific dummy variables, such as $GPT_Adoption_{i,t}$, $DeepSeek_Adoption_{i,t}$, $Gemini_Adoption_{i,t}$, which equal 1 if the firm has adopted the respective model by month t, based on explicit extractions from the prompt outputs. Compared to the more general adoption measures used in prior literature, this measure enables us to identify which specific GenAI model is being adopted by firms.

Figure 2 depicts the time series of GenAI adoption rates based on firms' responses on the IIPs from January 2022 to August 2025. The figure shows that the public rollouts

such as traditional machine learning, we exclude general terms like "AI" or "artificial intelligence" from our search query.

of ChatGPT and DeepSeek generate prominent adoption surges, underscoring their catalytic roles in accelerating practical integration of GenAI technologies among Chinese listed firms. Note that although OpenAI banned GPT accessibility from China, Microsoft Azure has been providing official and stable GPT access to enterprise customers.[4] The wide rollout of the service starts from Janurary 17th 2023,[5] which is also consistent with the first spike of GenAI Adoption in Figure 2. Table A1 summarizes the frequency with which specific GenAI models are mentioned in firms' responses. Consistent with Figure 2, GPT and DeepSeek dominate the discussion, appearing far more frequently than other models. These patterns validate our use of GPT and DeepSeek releases as benchmark milestones for foreign and domestic GenAI innovations.

[Insert Figure 2 here]

Figure 3 visualizes the semantic landscape of GenAI applications mentioned by firms. Specifically, it presents three-word clouds summarizing the most frequently cited application areas derived from firms' responses, including R&D and Technology Development, Operations and Management, and Products and Applications. Panel A highlights terms related to R&D and technological exploration, such as reasoning, multi-modality and knowledge; Panel B emphasizes operational and managerial applications, including management, workflow and decision-making; and Panel C focuses on product-level innovations and customer-facing uses, such as customer service, healthcare and education. These patterns reveal that firms deploy GenAI across a broad spectrum of functions, highlighting their relevance to firm value.

[Insert Figure 3 here]

2.2 Measure of Firm Sovereignty

To measure a firm's sovereignty exposure, we begin with four candidate proxies, each capturing the extent to which the firm aligns with China's national interests from different perspectives. We further extract their common component by estimating the first principal component across these proxies. This principal component serves as our primary measure of firm-level sovereignty exposure in the empirical analysis.

[4] See https://www.scmp.com/tech/big-tech/article/3283259/microsoft-closes-azure-subscription-individuals-access-openai-mainland-china

[5] https://azure.microsoft.com/en-us/blog/general-availability-of-azure-openai-service-expands-access-to-large-advanced-ai-models-with-added-enterprise-benefits

The first proxy is a binary indicator of state ownership (*SOE*), which equals 1 if the firm is ultimately controlled by a state entity and 0 otherwise. The decisions of state-owned enterprises (SOEs) often align with national development priorities and represent a nation's strategic interests (Bai, Lu and Tao, 2006). They thus capture the institutional dimension of national sovereignty.

The second proxy is an indicator for regulated industry (*RegInd*), which equals 1 if the firm operates in a sector the state designates as strategically important (e.g., energy, telecommunications, banking), and 0 otherwise. We define these sectors following the State Council's (2006) policy on safeguarding "the economic lifeline of the country," cross-referenced with Bradshaw et al. (2019). Firms in these industries play a vital role in maintaining economic security and national resilience.

The third proxy captures the degree of technological decoupling between Chinese and U.S. firms (*Decouple*). Following Han et al. (2024), we use patent citation data to construct a decoupling index that measures how frequently Chinese patents cite U.S. versus domestic patents. A lower share of citations to U.S. patents indicates less reliance on U.S. technology and thus corresponds to a higher value of the index. Firms with index values above the sample median are coded as *HighDecouple* = 1, and others as 0.

The fourth proxy captures a firm's exposure to U.S.-China tensions. Following Rogers et al. (2024), we estimate a firm's U.S.-China tension beta (*UCTBeta*) by regressing its monthly stock returns on changes in the U.S.-China tension index and the Fama-French three factors. Firms more exposed to geopolitical tensions tend to exhibit more negative *UCTBeta* values. We then construct a binary indicator of high detachment from the U.S., *HighUCTBeta*, which equals 1 if a firm's *UCTBeta* is above the sample median and 0 otherwise.

Finally, we construct an integrated measure of firm-level sovereignty exposure based on the four proxies using principal component analysis (PCA). To extract their common sovereignty component, we estimate the first principal component across the four proxies and use it as a unified measure of sovereignty exposure, denoted *SovIndex*. This measure combines standardized values of the four proxies as follows:

$$\begin{aligned} SovIndex = {} & 0.663 \times Z(SOE) + 0.648 \times Z(RegInd) \\ & + 0.306 \times Z(HighDecouple) + 0.215 \times Z(HighUCTBeta) \end{aligned}$$

The first principal component explains 32.9% of the total variance. We then construct a dummy variable for high sovereignty exposure, *HighSovIndex*, which equals one if a firm's *SovIndex* value is above the sample median and zero otherwise. This variable serves as the treatment indicator in our baseline regression.

We present the correlation matrix of the four proxies and the unified measure in the Panel B of Table 1. All proxies are positively and significantly correlated with one another at the 1% level. However, the magnitudes of these correlations are modest (ranging from 0.02 to 0.27), suggesting that the proxies capture complementary rather than overlapping dimensions of sovereignty. Specifically, *SOE* and the indicator of regulated industry (*RegInd*) identify firms strategically aligned with China's national interests, whereas the decoupling index (*HighDecouple*) and the U.S.-China tension exposure (*HighUCTBeta*) represent firms that are structurally decoupled from the United States. Collectively, these four proxies provide a comprehensive framework for quantifying firm-level sovereignty.

2.3 Sample Construction

Our sample comprises publicly listed Chinese firms that were questioned about generative AI (GenAI) on IIPs between August 2022 and August 2025. In total, 3656 out of 5260 listed firms received at least one GenAI-related inquiry during the sample period—representing 70% of all firms listed on the Chinese stock markets and 89% of market capitalization.[6] Firm financial characteristics and stock market variables are obtained from CSMAR and Wind, two widely used databases in empirical research on Chinese capital markets. The panel data comprises 153,598 firm–month observations. Panel A of Table 1 presents the summary statistics of the key variables used in this study. The average GenAI adoption rate is 0.169, with a standard deviation of 0.374, indicating substantial variation in GenAI adoption across firms.

Our empirical strategy exploits the staggered releases of two foundational milestone models—ChatGPT and DeepSeek R1—as shocks to firms' access to foreign and domestic AI systems, respectively. OpenAI released the ChatGPT Plus for commercial and retail use in February 2023, while DeepSeek R1 was officially

[6] As a preliminary validation test, we compare several alternative proxies for AI adoption between firms that receive GenAI-related questions and those that do not. The results in Table A2 show that firms receiving GenAI-related inquiries post more AI-related job openings, submit more AI-algorithm filings, and mention more AI-related terms in the MD&A sections of their annual reports. These patterns increase confidence that GenAI-related queries in the IIPs contain meaningful information about firms' AI-adoption behaviors.

launched and gained widespread attention in February 2025. For each shock, we define the corresponding release month as the event time and construct a 13-month event window (−6m, +6m) centered on it. The six-month pre- and post-event horizons allow sufficient time for observable adoption behavior to materialize, as new generative AI releases typically trigger immediate yet persistent adoption spikes (Fradkin, 2025). For the ChatGPT shock, we include all firms that recorded a GenAI question within the testing window (August 2022 to August 2023). For the DeepSeek shock, we restrict the sample to firms that were asked GenAI-related questions both before and after the DeepSeek release. This restriction mitigates measurement error by ensuring credible adoption information both before and after of the event.[7] There are 1,657 unique firms during the ChatGPT window and 1,507 firms during the DeepSeek window.

[Insert Table 1 here]

Panel C of Table 1 compares the firm characteristics between the GPT and DeepSeek samples. The two groups exhibit no material differences in key firm attributes. For example, the average log of total assets is 22.31 for the GPT sample and 22.43 for the DeepSeek sample, while the proportion of SOEs is 23.44% and 26.28%, respectively. These similarities help alleviate concerns that firms in the GPT sample differ systematically from those in the DeepSeek sample.

3. Validation of AI Adoption Measure

3.1 External Validation Using Established AI Adoption Proxies

We begin by performing a validation test of our GenAI adoption measure. We examine the relationship between our GenAI adoption measure and established AI adoption measures from the prior literature. If our measure effectively captures firms' AI adoption behavior, we would expect it to exhibit significant correlations with these existing proxies.

We employ three proxies for AI adoption following the literature (e.g., Webb, 2019; Babina et al., 2024; Eisfeldt et al., 2023). The first proxy is the number of AI-related terms appearing in the MD&A section of firms' annual reports ($AINumber_{i,t}$), following

[7] Our main findings remain robust under alternative sample selection criteria. First, we exclude the 2.4% of firms that did not respond to investors' questions. Second, we further exclude the 2.7% of firms whose responses contain no relevant information—typically referring investors to official documents or announcements without providing substantive answers. Third, we classify the 4.23% of firms that stated they "plan to adopt" GenAI models as non-adopters, and only those that indicated they "have adopted" as adopters. The results of these robustness tests are reported in Table A4 of the Internet Appendix.

the prior literature (e.g., Eisfeldt et al., 2023). Firms that reference AI-related content more frequently in their annual reports are more likely to be active AI adopters. The detailed procedure and full list of keywords are provided in Table A5 of the Appendix.

The second proxy is a binary variable equal to one if a firm files at least one internet information service algorithm with regulators ($I(Algorithm>0)_{i,t}$). Under Chinese regulations, providers of algorithmic recommendation services and artificial intelligence–generated content (AIGC) must complete a filing within ten working days of launch.[8] Because only firms with deployed algorithms are required to file, such filings provide an informative indicator of AI integration.

The third proxy is an indicator variable equal to one if a firm posts at least one AI-related job opening ($I(AIJobPosting>0)_{i,t}$) following the literature (e.g., Webb, 2019; Babina et al., 2024). We match job postings from major Chinese online job platforms to listed firms and identify AI-related positions using AI-related dictionary in job descriptions follows that of Babina et al. (2024). The AI-related keywords are provided in Table A5 of the Appendix.

[Insert Table 2 here]

Table 2 Panel A reports the relationship between our GenAI adoption measure and the three proxy variables over the 2022-2024 sample period. Because these proxies exhibit limited time-series variation within firms, we include year fixed effects and their lagged values in the regression. All three proxies are strongly and positively correlated with our GenAI adoption measure, with coefficients significant at the 1% level. As these proxies capture firms' engagement with GenAI, the positive correlations suggest that our IIPs-based measure serves as a reasonable indicator of firms' GenAI adoption behavior. Overall, these results support the validity of our IIP-based measure as a credible proxy for firm-level GenAI adoption.

3.2 Validation of LLM-Based Annotation Quality

[8] The regulations are the Provisions on the Administration of Algorithmic Recommendation for Internet Information Services, effective since March 2022, and the Provisions on the Administration of Deep Synthesis of Internet Information Services, adopted in November 2022. The filing requirement extends to a broad set of AIGC functions. These include text generation and conversational algorithms (e.g., GPT, Claude, DeepSeek), speech synthesis and transformation systems (e.g., Microsoft Azure TTS, WaveNet, Alibaba TTS), music and other non-speech audio generation (e.g., Google's MusicLM, Huawei's AudioTAG), biometric image and video editing such as face replacement (e.g., DeepMind's FaceNet, ByteDance's Faceu), non-biometric image and video generation or enhancement (e.g., Stable Diffusion, ERNIE-ViLG), as well as 3D reconstruction and virtual simulation (e.g., Omniverse, Tencent Digital Human).

To assess the reliability of using a large language model (LLM) rather than human experts for the classification task, we conduct a validation exercise on a random sample of 50 Q&As. While the efficiency advantages of LLM-based labeling are clear, the key question is whether the quality of LLM annotation is comparable to a human research assistant who could feasibly be hired for this task.

We benchmark LLM-generated labels to manual annotations and find a high degree of agreement, as shown in Panel B of Table 2. Across the sample, the LLM achieves an accuracy of 92%, indicating very high agreement between human and machine labels. The level of accuracy is comparable to that reported in prior work, such as Eisfeldt et al. (2023), and suggests that LLM-generated classifications are sufficiently reliable for large-scale empirical analysis.

4. Sovereignty and GenAI Adoption

In this section, we empirically examine the hypotheses on AI sovereignty. We begin by estimating the baseline specification using each of the four sovereignty-exposure proxies individually as a preliminary analysis. We then employ the PCA-based unified index as our primary measure of sovereignty exposure to re-estimate the baseline regression and conduct the subsequent empirical analyses.

As a preliminary illustration, in Figure 4 we plot time-series data for AI adoption rates between sovereignty-exposed and less-sovereignty-exposed firms. The graphs include all firms questioned by investors via IIPs during the sample period. Across the four sovereignty proxies—*SOE*, *RegInd*, *HighDecouple*, and *HighUCTBeta*—a consistent pattern emerges. Following the release of ChatGPT, sovereignty-exposed firms exhibit substantially lower AI adoption rates compared to their less sovereignty-exposed counterparts. However, this pattern reverses after the release of DeepSeek, with the adoption gap narrowing significantly. This figure offers preliminary support for the main finding of our study. Firms more closely aligned with national interests, or those more detached from the United States, initially hesitate to adopt foreign AI technologies. Once competitive domestic models become available, these firms accelerate their adoption.

[Insert Figure 4 here]

Figure 5 further illustrates the cross-industry heterogeneity in early adoption and subsequent catch-up. Industries such as information services, professional and business

services, and trade contain a larger share of early adopters, especially for firms with lower sovereignty exposure. In contrast, sectors like financials and real estate, transportation, agriculture and extraction exhibit delayed adoption, catching up only after the release of DeepSeek, and are characterized by higher sovereignty exposure.

[Insert Figure 5 here]

4.1 Four Sovereignty Proxies and GenAI Adoption

We formally examine the impact of sovereignty exposure on GenAI adoption within an event-study framework. We estimate the following baseline specification separately for the ChatGPT and DeepSeek releases:

$$AI_Adoption_{i,t} = \alpha + \beta \times SovProxy_i \times PostAI_t + \gamma_i + \delta_t + \varepsilon_{i,t}$$

where $AI_Adoption_{i,t}$ is a dummy variable equal to 1 if firm i has indicated GenAI adoption in its responses on the Investor Interaction Platforms by month t. $SovProxy_i$ represents one of four sovereignty proxies—*SOE*, *RegInd*, *HighDecouple*, or *HighUCTBeta*. $PostAI_t$ is a post-event indicator that equals 1 in months following the corresponding AI release. All regressions include firm fixed effects (γ_i) and year–month fixed effects (δ_t).

For the GPT sample, the coefficient β captures how sovereignty exposure influences firms' adoption decisions following the initial foreign GenAI release. For the DeepSeek sample, because $AI_Adoption_{i,t}$ is an absorbing outcome (once equal to one, it remains one), the identification relies exclusively on firms that had not adopted prior to January 2025. Firms that had already adopted before the DeepSeek release provide no within-firm variation and are thus absorbed by the firm fixed effects. As a result, β measures the extent to which sovereignty-exposed firms transition from non-adopters to adopters after the release of DeepSeek.

[Insert Table 3 here]

Panel A of Table 3 reports the results for the ChatGPT shock. Across all specifications, the coefficients on *SovProxy*×*PostGPT* are negative, with three of the four estimates statistically significant at the 5% or 1% level. The effects are most pronounced when sovereignty exposure is measured by technology-decoupling or U.S.–China tension. These results indicate that firms with greater sovereignty exposure are significantly less likely to adopt GenAI following the release of ChatGPT.

Panel B reports the results for the DeepSeek shock. In contrast to Panel A, the coefficients on *SovProxy*×*PostDS* are positive and statistically significant across all four proxies, with magnitudes comparable to those in Panel A. These positive coefficients indicate a reversal in GenAI adoption behavior: firms more exposed to sovereignty concerns—previously lagging in adopting foreign models—accelerated their adoption once high-quality domestic GenAI systems became available.

Panel C reports a horse-race specification that includes all four sovereignty proxies in a single regression. For the ChatGPT shock, the coefficients on the interactions of *PostGPT* with *SOE*, *HighDecouple*, and *HighUCTBeta* are significantly negative. For the DeepSeek shock, the interactions of *PostDS* with *SOE* and *HighDecouple* remain significantly positive. These results suggest that the four proxies capture complementary dimensions of sovereignty that collectively explain firms' divergent responses to foreign versus domestic AI shocks.

Taken together, the results show that sovereignty-exposed firms delay the adoption of foreign GenAI but rapidly embrace GenAI once quality domestic models emerge.

4.2 The PCA-based Sovereignty Index and GenAI Adoption

We next employ the PCA-based unified measure of sovereignty, *HighSovIndex*, to re-estimate the baseline regression and conduct the subsequent empirical analyses. To isolate variation net of industry- and region-specific shocks, we further include stringent time–industry and time–city fixed effects.

[Insert Table 4 and Figure 6 here]

Table 4 reports the results. Columns (1) and (2) examine the impact of the ChatGPT release. The negative coefficients on the interaction term, *HighSovIndex*×*PostGPT*, indicate that firms with higher sovereignty exposure experienced significantly lower AI adoption following the ChatGPT shock compared to low-exposure firms. Economically, the coefficient of -0.062 (t = -2.19) implies that high-exposure firms exhibited a 17.7% smaller increase in GenAI adoption, using the standard deviation of GenAI adoption as the benchmark. In contrast, columns (3) and (4) show the results for the DeepSeek release. The interaction coefficient becomes positive (0.063, t = 2.38), suggesting that highly exposed firms experienced a 17.9% larger increase in GenAI adoption relative to low-exposure firms. These findings are highly consistent with the earlier results based on the individual sovereignty proxies.

A key advantage of our IIPs-based measure is its ability to identify which specific model firms adopt. To enhance identification power, we replace the general GenAI adoption measure with indicators of model-specific adoption—namely, the adoption of ChatGPT (*GPT_Adoption*) and DeepSeek (*DeepSeek_Adoption*). The results, reported in Table A3 of the Internet Appendix, are consistent with our baseline findings: firms with higher sovereignty exposure are significantly less likely to adopt ChatGPT following the GPT release, but more likely to adopt DeepSeek after its launch. By explicitly matching each model's release event to the adoption of the same model, the analysis sharpens identification and reinforces the validity of the results.

We further examine the parallel trends assumption by plotting dynamic treatment effects around each release event in Figure 6. Panel A shows no significant difference in GenAI adoption trends between high- and low-exposure firms before the ChatGPT release, followed by a sharp negative divergence afterward. Similarly, Panel B reveals no difference in GenAI adoption prior to the DeepSeek release, but a positive and significant divergence afterward. These patterns support the parallel trends assumption.

In summary, the PCA-based results are consistent with the notion of safeguarding AI sovereignty: firms more exposed to sovereignty concerns initially lagged in adopting foreign AI technologies but subsequently caught up once competitive domestic alternatives became available.

4.3 Alternative explanations

GPT and DeepSeek differ along several dimensions beyond nationality. It could be the differences unrelated to nationality that lead to the observed GenAI adoption behaviors. To assess this possibility, we examine two alternative explanations: the licensing model of the software (open-source vs. closed-source) and the adoption costs associated with each GenAI model.

4.3.1 Open-source vs Closed-source

A key difference between GPT and DeepSeek lies in their licensing model: GPT is primarily closed-source, whereas DeepSeek is open-source. Closed-source models typically require external deployment, limiting data privacy and restricting fine-tuning for customized applications. Open-source models, by contrast, can often be deployed locally, enabling greater control over data and allowing firms to train or fine-tune the

model to meet specific needs. These characteristics may influence adoption decisions, with firms prioritizing data privacy or customization more likely to adopt open-source solutions. Therefore, one alternative explanation is that the observed adoption behavior reflects not a nationalism preference, but rather firms' differing attitudes toward open- (DeepSeek) versus closed-source models (ChatGPT).

[Insert Table 5 here]

To test this, we examine firms' responses to the releases of models that share the same nationality but differ in licensing: Gemini (foreign, closed-source) and LLaMA (foreign, open-source). If the preference were made on the basis of whether a model was open-source-licensed, then we would expect that sovereignty-exposed firms would prefer LLaMA to Gemini. However, as shown in Table 5, sovereignty-exposed firms exhibit consistently negative adoption responses to both Gemini and LLaMA (β = -0.002, $p < 0.05$). This pattern contradicts the open- vs. closed-source interpretation and instead supports the sovereignty-based interpretation.

4.3.2 The Low-Cost Hypothesis

Another salient feature of DeepSeek is its substantially lower cost relative to ChatGPT. For example, in January2025, the per-one million tokens price for DeepSeek was $1.1, compared to $20 for the sate-of-art GPT counterpart. It is therefore possible that firms with sovereignty exposure prefer DeepSeek mainly for its cost advantage rather than its domestic origin. To examine this possibility, we include proxies for firms' cost sensitivity and financial constraints, along with their interactions with the post-AI indicator, in the baseline regression. If the pecuniary cost of using GenAI models drives firms' choices, we would expect the baseline effect to attenuate or even vanish after controlling for these cost-related firm characteristics.

Specifically, we measure cost sensitivity with markup, defined as the percentage difference of output prices and marginal cost. A lower markup means the firm is less likely to transmit cost shocks to the downstream and is thus more cost sensitive. Without firm-level quantity data, we proxy for markup with the percentage difference of gross revenue and the cost of goods sold. We then use the sample median of this ratio to construct a dummy variable, *CostSensitive*, which equals one if the ratio is below the sample median and zero otherwise. To capture firms' financial constraints, we employ the SA index (Hadlock and Pierce, 2010), the KZ index (Kaplan and Zingales, 1997),

and the WW index (Whited and Wu, 2006). We then examine whether incorporating these proxies alters the baseline results.

[Insert Table 6 here]

Panels A and B of Table 6 present the results. Across all specifications, the coefficients on *HighSovIndex*×*PostGPT* remain negative and significant, while those on *HighSovIndex*×*PostDS* remain positive and significant. The economic magnitudes of these coefficients are also comparable to the baseline estimates reported in Table 4. In contrast, the coefficients on the interaction terms between the post-AI indicator and the proxies for cost sensitivity and financial constraints are generally insignificant. The evidence suggests that differences in the adoption costs of GenAI models are unlikely to be the main driver of the divergent behaviors observed among firms with varying degrees of exposure to AI sovereignty.

4.4. Additional Analysis: State vs. Market Trade-Off

We next conduct two additional analyses to examine the underlying forces that shape firms' decisions to adopt GenAI technologies. From a political economy perspective, firms' choice of GenAI models embodies a classic "state versus market" trade-off (e.g., Dixit et al., 1997; Bai et al., 2006; Ljungqvist et al., 2015). On one hand, the state has strong imperatives to safeguard technological sovereignty and protect national strategic interests. These objectives may incentivize firms to align their technology adoption choices with national priorities, to avoid political or regulatory costs or to secure preferential treatment from the state. On the other hand, foreign GenAI models may offer superior and often more timely economic value enhancement compared to domestic alternatives—by improving efficiency, stimulating innovation, and improving profitability. Such economic advantages may lead firms to favor foreign GenAI models, despite potential political or institutional pressures to support domestic technologies.

A firm's ultimate choice thus depends on the relative strength of state forces versus market forces. Because these forces vary across regions and industries, we expect firms' preferences for GenAI models to differ accordingly. To test this intuition, we exploit cross-sectional variation in both state power and market competition to examine how these factors shape firms' GenAI adoption behaviors. Intuitively, if firms' adoption decisions are indeed influenced by sovereignty considerations, we should observe

stronger baseline effects in regions characterized by greater state influence and weaker market competition.

4.4.1 State Influence

We begin by examining whether the preference for domestic GenAI is more pronounced in regions with stronger state influence. We employ three proxies to capture government influence: (i) the ratio of local government expenditure to regional GDP, (ii) geographical proximity to political centers (i.e., Beijing versus non-Beijing regions), and (iii) the presence of political connections (i.e., whether senior management has previously held government positions). A higher ratio of government spending to GDP indicates greater dependence of the local economy on government activity, and hence stronger government influence on local firms. Proximity to the political center is inversely related to the monitoring costs of the central government and thus often serves as an additional proxy for government influence (Duchin et al., 2020). Similarly, firms with politically connected executives are more likely to be subject to government influence (Faccio, 2006; Goldman et al., 2009).

[Insert Table 7 here]

Table 7 reports the results. Following the ChatGPT shock, sovereignty-exposed firms located in high–government-spending regions, in Beijing, and those with politically connected executives exhibit significantly lower adoption of foreign AI technologies. For instance, the coefficient on *HighSovIndex*×*PostGPT* is -0.169 (t =-2.00) for firms headquartered in Beijing, compared with -0.041 (t=-1.34) for firms outside Beijing. The difference is statistically significant at the 1% level.

In contrast, after the DeepSeek shock, the same groups demonstrate significantly higher adoption of domestic GenAI. The coefficient on *HighSovIndex*×*PostDS* is 0.111 ($p < 0.01$) in high–government-spending regions and 0.115 for Beijing firms, with differences between the high and low groups statistically significant at the 1–5% levels. Collectively, these findings suggest that the defense of AI sovereignty is more pronounced where state influence is stronger.

4.4.2 Market Competitiveness

Next, we examine whether market competition affects firms' GenAI adoption decisions. Firms operating in less competitive industries face weaker market discipline,

granting them greater flexibility to align their decisions with state interests and potentially capture long-term benefits. In contrast, firms in highly competitive industries may prioritize efficiency over geopolitical alignment due to intensified short-term market pressures. We measure market competition using the Herfindahl–Hirschman Index (HHI), defining firms in the top tercile of HHI as the low-competition group and those in the bottom tercile as the high-competition group. We re-estimate the baseline regression separately for these two groups, with results reported in Table 8.

[Insert Table 8 Here]

Following the ChatGPT shock, high-sovereignty firms in concentrated industries exhibit significantly stronger reluctance toward foreign GenAI (β=-0.076, $p<0.10$) compared with low-sovereignty firms, whereas this difference is insignificant in competitive industries. Similarly, after the release of DeepSeek, high-sovereignty firms in concentrated industries adopt domestic GenAI more actively (β=0.099, $p<0.10$) than their low-sovereignty counterparts, while no significant difference is observed in highly competitive industries. Collectively, these results suggest that high market competition attenuates the sovereignty-driven reluctance toward foreign GenAI and preference for domestic GenAI.

4.5 Drivers of the Defense of AI Sovereignty: Data Security Concerns

A potential factor underlying sovereignty-exposed firms' reluctance to adopt foreign GenAI systems is concern over data security. Data are a fundamental resource for training advanced models and have gained heightened strategic importance in the era of AI (e.g., Farboodi and Veldkamp, 2021; Veldkamp and Chung, 2024; Abis and Veldkamp, 2023). For example, China projects that its investment in data infrastructure will exceed USD 280 billion between 2025 and 2030.[9]

However, data security risks associated with GenAI applications have also become increasingly prominent[10]. Data leakage through AI systems can undermine a country's technological and informational advantages and threaten its national security (e.g., Chang et al., 2023; Sun and Trefler, 2023). These concerns are especially pronounced for high-sovereignty firms, given their greater exposure to national interests and

[9] Source: https://www.news.cn/tech/20250107/85bd4e962d17436c9d80a5a690f77586/c.html

[10] For example, according to IBM, 96% of executives say adopting GenAI makes a security breach likely in their organizations in the next three years. https://www.ibm.com/think/news/are-new-genai-tools-putting-your-business-at-risk

security issues in the GenAI race. To mitigate these risks, China has enacted comprehensive legal and administrative measures. A series of national laws—Cybersecurity Law (2017), Data Security Law (2021), and Personal Information Protection Law (2021)—establish stringent rules governing data processing, cross-border data transfers, and algorithmic governance. Accordingly, high-sovereignty firms in China are likely to face heightened compliance and legal risks when adopting generative AI models. These risks, in turn, create strong incentives to favor domestic GenAI models over foreign ones.

4.5.1 Data Security Concerns: The Impact on the Adoption Decisions

To test this prediction, we construct two proxies for firms' exposure to data security risks. The first proxy measures the frequency of data asset-related terms appearing in the Management Discussion and Analysis (MD&A) sections of annual reports. A higher frequency of such terms indicates a stronger linkage between firm value and data assets, and hence greater exposure to data security risks. We begin with four seed words—*Information*, *Network*, *Digital*, and *Data*—and expand them into 137 data asset-related keywords using a Word2Vec model (Mikolov et al., 2013). Details of this procedure and the full keyword list are provided in Table A5 of the Appendix.

The second proxy captures firms' legal exposure to data breaches, as inferred from their annual reports. We first extract legal definitions of cross-border data breach activities from relevant Chinese data security laws.[11] These definitions are then incorporated into a structured prompt that instructs DeepSeek to analyze firms' reports and assess their data breach legal risks. This proxy reflects firms' potential legal and compliance risks associated with cross-border data leakage when utilizing foreign GenAI systems. Details of this process are provided in the Appendix.

[Insert Table 9 here]

Based on these two proxies, we classify sample firms into high and low data risk exposure groups and re-estimate the baseline regressions separately for each group. Table 9 shows that firms' reluctance to adopt foreign GenAI models and their preference for domestic ones are consistently stronger among firms with high data risk exposure. For instance, in Panel A, the coefficients on the interaction terms between

[11] These laws are *Data Security Law*, *Personal Information Protection Law*, *Regulation on Critical Information Infrastructure Security Protection*, and *Measures for Data Export Security Assessment*.

HighSovIndex and *PostAI* indicators are statistically significant only for the high data risk group but insignificant for the low exposure group. The difference in coefficients between the two groups is also significant at the 5% level. Similar results are observed in Panel B, where we use the median of data breach risk to split the sample. Overall, the evidence suggests that data security concerns are a major factor driving the sovereignty-induced reluctance toward foreign models and the preference for domestic ones.

4.5.2 Data Security Concerns: The Impact on the Deployment Choices

Table 9 examines the effect of data security on firms' *adoption decisions*: firms with greater data security concerns are more likely to prefer domestic GenAI models over foreign ones. To complement this analysis, we further examine the effect on *deployment choices*. Specifically, among firms that have adopted foreign GenAI models, we expect those with higher sovereignty exposure to apply these models in areas with lower legal and data-related risks. In contrast, for firms adopting domestic GenAI models, such an effect is not expected, as concerns over data security are substantially lower for these models.

To test this prediction, we analyze the relationship between firms' sovereignty exposure and the risk levels of the areas in which they apply GenAI. Firms often disclose how they use GenAI in their responses to inquiries on IIPs. For example, some firms report using GenAI in customer service, whereas others apply it in R&D activities. These applications vary in their data security risks—customer service represents a low-risk area, while R&D involves higher data sensitivity. We employ a prompt-based approach using DeepSeek V3 to evaluate firms' responses and assign each firm a data security risk score based on the nature of its GenAI usage. We then examine the relationship between this risk score and firms' sovereignty exposures separately for firms adopting GPT and those adopting DeepSeek.

[Insert Table 10 here]

The results are reported in Table 10. In column (1), where the sample includes firms adopting GPT, the coefficient on the high-sovereignty-exposure indicator is negative and highly significant at the 1% level. In contrast, the corresponding coefficient in column (2), which includes firms adopting DeepSeek, is statistically insignificant ($t = 0.33$). These results indicate that firms with high sovereignty exposure

are more likely to use GPT in low–data-risk areas, consistent with strategic risk management behavior for foreign GenAI models. By contrast, firms adopting DeepSeek show no such pattern, likely because domestic models pose fewer data security concerns.

Figures 7 and 8 provide visual evidence consistent with these regression results. Figure 7 contrasts the keywords of GenAI applications between high- and low-data-risk applications, showing that high-risk cases emphasize finance, healthcare and security whereas low-risk cases feature customer service, computing and generation. Figure 8 further compares GPT and DeepSeek adopters: GPT-related keywords closely resemble those in the low-risk group, while DeepSeek-related terms align more with high-risk activities.

Together, these patterns suggest that firms strategically allocate foreign GenAI models to safer, less data-sensitive domains, while domestic models are more integrated into core, data-intensive functions.

[Insert Figures 7 and 8 here]

4.6 Market Reactions

How the stock market perceives high sovereignty firms' GenAI adoption also offers valuable insight into the motivations behind these decisions. Foreign GenAI products such as ChatGPT are generally regarded as more advanced than China's domestic counterparts, such as DeepSeek. A preference for domestic GenAI may therefore place Chinese firms at a competitive disadvantage relative to foreign peers, potentially leading to a negative stock market response. However, if safeguarding AI sovereignty is viewed as critical to maintaining China's long-term position in the global GenAI race, reliance on domestic models may generate long-run value, particularly for firms with greater exposures to AI sovereignty concerns. Therefore, how the stock market evaluates firms' AI adoption decisions remains an open empirical question.

We first examine the average stock market reactions to firms' disclosures of GenAI adoption on IIPs. If our measure effectively captures information on GenAI adoption at a daily frequency, we should observe significant market responses around these disclosure events. For each adopting firm, we compute the cumulative abnormal returns (CARs) around its first explicit disclosure of GenAI adoption. We then regress these CARs on an indicator for overall GenAI adoption (*AI_Adoption*).

The results are reported in Panel A of Table 11. Across the CAR specifications, the coefficients on the general GenAI adoption indicator range from 0.005 to 0.07, all statistically significant at the 1% level. These estimates indicate that, on average, firms' announcements of GenAI adoption receive positive market reactions. This finding is consistent with prior studies documenting that GenAI adoption enhances firms' long-term fundamentals (Babina et al., 2025; Eisfeldt et al., 2023). Moreover, the significant market reactions at the daily frequency further highlight our measure's ability to capture high-frequency, value-relevant information on GenAI adoption.

[Insert Table 11 here]

We next examine whether the stock market reacts differently to GenAI adoption by high- versus low-sovereignty firms. For each firm disclosing its adoption of GenAI, we compute cumulative abnormal returns (CARs) around the disclosure date. Specifically, we estimate the following regression model:

$$CAR[-1, +1]_{i,t} = \alpha + \beta \times HighSovIndex_i \times AdoptDomAI_{i,t} + \gamma_1 \times HighSovIndex_i + \gamma_2 \times AdoptDomAI_{i,t} + \delta_t + \varepsilon_{i,t}$$

where $AdoptDomAI_{i,t}$ is a dummy variable equal to 1 if firm i adopts DeepSeek, and 0 if it adopts GPT. δ_t represents time fixed effects, including year-month and weekday fixed effects. We do not include firm fixed effects because this is a cross-sectional regression.

The results reported in Panel B of Table 11 are consistent with the earlier findings. In column (1), where CAR is based on CAPM, the coefficient on *HighSovIndex*×*AdoptDomAI* is 0.012 and statistically significant at the 1% level (t = 2.96). This positive coefficient indicates that, upon announcing the adoption of domestic (foreign) GenAI, the three-day CAR for high-sovereignty firms is 1.2 percentage points higher (lower) than that for low-sovereignty firms. Using low-sovereignty as the benchmark, the evidence suggests that the market responds positively to high-sovereignty firms' adoption of domestic GenAI, but does not view their adoption of foreign GenAI as value-enhancing.

5. Conclusion

This paper provides evidence that the global diffusion of generative AI is not a purely economic process but one deeply intertwined with national strategic priorities. We introduce the concept of AI sovereignty and show that firms exposed to national

interests and geopolitical decoupling, which we term as sovereignty-exposed firms, act as key agents in defending national technological sovereignty. These firms have a clear and systematic preference for domestically developed GenAI models rather than competing foreign ones. Using novel high-frequency data based upon investor-firm dialogues, we document that sovereignty-exposed firms were initially reluctant to adopt ChatGPT but became active adopters after the release of China's DeepSeek, domestic counterpart to ChatGPT.

Our results reveal that this sovereignty-driven behavior is heightened where state influence is stronger, and lessened where market competition is stronger. We find further that the underlying mechanism pointing to greater concern over data security. Moreover, equity market responses show that investors reward firms that align with these sovereignty considerations, indicating that market participants increasingly understand the strategic value of domestic AI alignment in a bifurcating global technology landscape.

These findings imply that the direction of frontier technologies such as GenAI is deeply influenced by competing national powers and can be increasingly segmented. As countries race for leadership in the artificial generative intelligence (AGI), firms exposed to sovereignty issues are likely to be pivotal in the implementation of national interests, shaping domestic innovation ecosystems, and redefining the global contours of technological rivalry.

Given the data–AI reinforcement loop, sovereignty-induced segmentation potentially set in motion a self-reinforcing dynamic: firms' model choices shape the data available to models, which in turn shapes subsequent training and innovation. This feedback loop may drive a growing divergence in the *path of technological evolution*, with national data regimes and AI architectures increasingly defining separate frontiers of progress. We leave the exploration of this long-run divergence mechanism to future research.

References

Abis, Simona, and Laura Veldkamp. 2023. "The Changing Economics of Knowledge Production." *The Review of Financial Studies* 37 (1): 89–118. https://doi.org/10.1093/rfs/hhad059.

Acemoglu, Daron. 2025. "Nobel Lecture: Institutions, Technology, and Prosperity." *American Economic Review* 115 (6): 1709–48. https://doi.org/10.1257/aer.115.6.1709.

Agrawal, Ajay, Joshua Gans, Avi Goldfarb, and Catherine Tucker. 2025. *The Political Economy of Artificial Intelligence*. University of Chicago Press. https://www.nber.org/books-and-chapters/political-economy-artificial-intelligence.

Babina, Tania, Anastassia Fedyk, Alex He, and James Hodson. 2024. "Artificial Intelligence, Firm Growth, and Product Innovation." *Journal of Financial Economics* 151 (January): 103745. https://doi.org/10.1016/j.jfineco.2023.103745.

Bai, Chong-En, Jiangyong Lu, and Zhigang Tao. 2006. "The Multitask Theory of State Enterprise Reform: Empirical Evidence from China." *American Economic Review* 96 (2): 353–57. https://doi.org/10.1257/000282806777212125.

Benguria, Felipe, and Felipe Saffie. 2021. "Beyond Tariffs: How Did China's State–Owned Enterprises Shape the US–China Trade War?" SSRN Scholarly Paper 3969215. Social Science Research Network, November 15. https://doi.org/10.2139/ssrn.3969215.

Beraja, Martin, Andrew Kao, David Y Yang, and Noam Yuchtman. 2023. "AI-Tocracy." *The Quarterly Journal of Economics* 138 (3): 1349–402. https://doi.org/10.1093/qje/qjad012.

Beraja, Martin, Wenwei Peng, David Yang, and Noam Yuchtman. 2024. *Government as Venture Capitalists in AI*. W32701. National Bureau of Economic Research. https://doi.org/10.3386/w32701.

Beraja, Martin, David Y Yang, and Noam Yuchtman. 2023. "Data-Intensive Innovation and the State: Evidence from AI Firms in China." *The Review of Economic Studies* 90 (4): 1701–23. https://doi.org/10.1093/restud/rdac056.

Bick, Alexander, Adam Blandin, and David J. Deming. 2024. "The Rapid Adoption of Generative AI." Working Paper 32966. Working Paper Series. National Bureau of Economic Research, September. https://doi.org/10.3386/w32966.

Bonney, Kathryn, Cory Breaux, Catherine Buffington, et al. 2024. "Tracking Firm Use of Ai in Real Time: A Snapshot from the Business Trends and Outlook Survey." SSRN Scholarly Paper 4794368. Social Science Research Network, April 1. https://papers.ssrn.com/abstract=4794368.

Bradshaw, Mark, Guanmin Liao, and Mark (Shuai) Ma. 2019. "Agency Costs and Tax Planning When the Government Is a Major Shareholder." *Journal of Accounting and Economics* 67 (2–3): 255–77. https://doi.org/10.1016/j.jacceco.2018.10.002.

Brynjolfsson, Erik, Anton Korinek, and Ajay Agrawal. 2025. *A Research Agenda for the Economics of Transformative AI*. W34256. National Bureau of Economic Research. https://doi.org/10.3386/w34256.

Canayaz, Mehmet, Isil Erel, Umit Gurun, and Yufeng Wu. 2024. *When Protectionism Kills Talent*. W32466. National Bureau of Economic Research. https://doi.org/10.3386/w32466.

Cen, Ling, Lauren Cohen, Jing Wu, and Fan Zhang. 2024. *Who Benefits from Trade Wars?* W32621. National Bureau of Economic Research. https://doi.org/10.3386/w32621.

Cen, Xiao, Vyacheslav Fos, and Wei Jiang. 2024. "How Do U.S. Firms Withstand Foreign Industrial Policies?" Working Paper 32411. Working Paper Series. National Bureau of Economic Research, May. https://doi.org/10.3386/w32411.

Chang, Qing, Lin William Cong, Liyong Wang, and Longtian Zhang. 2023. *Production, Trade, and Cross-Border Data Flows*. W31416. National Bureau of Economic Research. https://doi.org/10.3386/w31416.

Chen, Tuo, Chang-Tai Hsieh, and Zheng Michael Song. 2022. *Non-Tariff Barriers in the U.S.-China Trade War*. W30318. National Bureau of Economic Research. https://doi.org/10.3386/w30318.

Chen, Wilbur, and Suraj Srinivasan. 2024. "Going Digital: Implications for Firm Value and Performance." *Review of Accounting Studies* 29 (2): 1619–65. https://doi.org/10.1007/s11142-023-09753-0.

Clayton, Christopher, Antonio Coppola, Matteo Maggiori, and Jesse Schreger. 2025. "Geoeconomic Pressure." Working Paper 34020. Working Paper Series. National Bureau of Economic Research, July. https://doi.org/10.3386/w34020.

Clayton, Christopher, and Matteo Maggiori. 2025. "A FRAMEWORK FOR GEOECONOMICS." *Econometrica*. *Forthcoming*

Clayton, Christopher, Matteo Maggiori, and Jesse Schreger. 2025. "The Political Economy of Geoeconomic Power." Working Paper 33353. Working Paper Series. National Bureau of Economic Research, January. https://doi.org/10.3386/w33353.

Crosignani, Matteo, Lina Han, Marco Macchiavelli, and André F. Silva. 2024. *Geopolitical Risk and Decoupling: Evidence from U.S. Export Controls*. Staff Reports (Federal Reserve Bank of New York). Staff Reports (Federal Reserve Bank of New York). Federal Reserve Bank of New York. https://doi.org/10.59576/sr.1096.

Dixit, Avinash. 1997. "Power of Incentives in Private versus Public Organizations." *The American Economic Review* 87 (2): 378–82.

Duchin, Ran, Zhenyu Gao, and Haibing Shu. 2020. "The Role of Government in Firm Outcomes." *The Review of Financial Studies* 33 (12): 5555–93. https://doi.org/10.1093/rfs/hhaa005.

Eisfeldt, Andrea L., Gregor Schubert, Miao Ben Zhang, and Bledi Taska. 2023. "Generative AI and Firm Values." SSRN Scholarly Paper 4436627. Social Science Research Network, May 2. https://doi.org/10.2139/ssrn.4436627.

Faccio, Mara. 2006. "Politically Connected Firms." *American Economic Review* 96 (1): 369–86. https://doi.org/10.1257/000282806776157704.

Farboodi, Maryam, and Laura Veldkamp. 2021. *A Model of the Data Economy*. W28427. National Bureau of Economic Research. https://doi.org/10.3386/w28427.

Fisman, Raymond, Yasushi Hamao, and Yongxiang Wang. 2014. "Nationalism and Economic Exchange: Evidence from Shocks to Sino-Japanese Relations." *Review of Financial Studies* 27 (9): 2626–60. https://doi.org/10.1093/rfs/hhu017.

Fradkin, Andrey. n.d. *Demand for LLMs: Descriptive Evidence on Substitution, Market Expansion, and Multi-Homing*. Working Paper

Goldman, Eitan, Jörg Rocholl, and Jongil So. 2009. "Do Politically Connected Boards Affect Firm Value?" *Review of Financial Studies* 22 (6): 2331–60. https://doi.org/10.1093/rfs/hhn088.

Hadlock, Charles J., and Joshua R. Pierce. 2010. "New Evidence on Measuring Financial Constraints: Moving Beyond the KZ Index." *The Review of Financial Studies* 23 (5): 1909–40. https://doi.org/10.1093/rfs/hhq009.

Han, Pengfei, Wei Jiang, and Danqing Mei. 2024. "Mapping U.S.–China Technology Decoupling: Policies, Innovation, and Firm Performance." *Management Science* 70 (12): 8386–413. https://doi.org/10.1287/mnsc.2022.02057.

Jha, Manish, Jialin Qian, Michael Weber, and Baozhong Yang. 2025. "ChatGPT and Corporate Policies." SSRN Scholarly Paper 4521096. Social Science Research Network, February 28. https://doi.org/10.2139/ssrn.4521096.

Kaplan, Steven N., and Luigi Zingales. 1997. "Do Investment-Cash Flow Sensitivities Provide Useful Measures of Financing Constraints?." *The Quarterly Journal of Economics* 112 (1): 169–215. https://doi.org/10.1162/003355397555163.

Lee, Charles M. C., and Qinlin Zhong. 2022. "Shall We Talk? The Role of Interactive Investor Platforms in Corporate Communication." *Journal of Accounting and Economics* 74 (2): 101524. https://doi.org/10.1016/j.jacceco.2022.101524.

Ljungqvist, Alexander, Donghua Chen, Dequan Jiang, Haitian Lu, and Mingming Zhou. 2015. *State Capitalism vs. Private Enterprise*. W20930. National Bureau of Economic Research. https://doi.org/10.3386/w20930.

Mikolov, Tomas, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. "Efficient Estimation of Word Representations in Vector Space." arXiv:1301.3781. Preprint, arXiv, September 10. https://doi.org/10.48550/arXiv.1301.3781.

Petrova, Maria, Gregor Schubert, Bledi Taska, and Pinar Yildirim. 2024. *Career Values for Labor Markets: Evidence from Robot Adoption*. W32655. National Bureau of Economic Research. https://doi.org/10.3386/w32655.

Rogers, John H., Bo Sun, and Tony Sun. 2024. "U.S.-China Tension." *SSRN Electronic Journal*, ahead of print. https://doi.org/10.2139/ssrn.4815838.

Schubert, Gregor. 2025. "Organizational Technology Ladders: Remote Work and Generative AI Adoption." Preprint, Elsevier BV. https://doi.org/10.2139/ssrn.5094265.

Sheng, Jinfei, Zheng Sun, Baozhong Yang, and Alan L. Zhang. 2024. "Generative AI and Asset Management." SSRN Scholarly Paper 4786575. Social Science Research Network, August 1. https://papers.ssrn.com/abstract=4786575.

Sun, Ruiqi, and Daniel Trefler. 2023. "The Impact of AI and Cross-Border Data Regulation on International Trade in Digital Services: A Large Language Model." Working Paper 31925. Working Paper Series. National Bureau of Economic Research, November. https://doi.org/10.3386/w31925.

Veldkamp, Laura, and Cindy Chung. 2024. "Data and the Aggregate Economy." *Journal of Economic Literature* 62 (2): 458–84. https://doi.org/10.1257/jel.20221580.

Webb, Michael. 2019. "The Impact of Artificial Intelligence on the Labor Market." SSRN Scholarly Paper 3482150. Rochester, NY, November 6. https://doi.org/10.2139/ssrn.3482150.

Whited, Toni M., and Guojun Wu. 2006. "Financial Constraints Risk." *The Review of Financial Studies* 19 (2): 531–59. https://doi.org/10.1093/rfs/hhj012.

Wong, T.J., Gwen Yu, Shubo Zhang, and Tianyu Zhang. 2024. "Calling for Transparency: Evidence from a Field Experiment." *Journal of Accounting and Economics* 77 (1): 101604. https://doi.org/10.1016/j.jacceco.2023.101604.

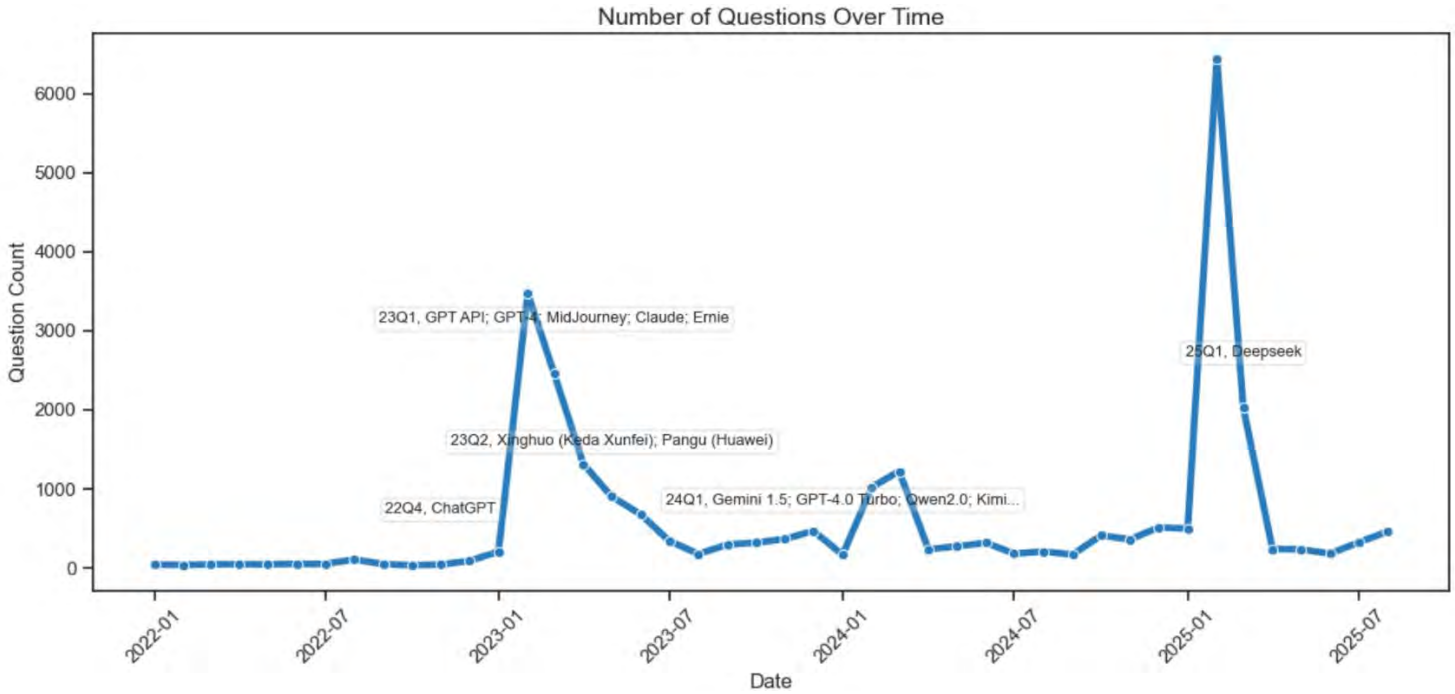


Figure 1. Number of Questions Over Time

This figure shows the time series of the monthly number of investor questions containing selected generative AI-related keywords, as recorded on the investor interaction platforms. The sample period is from 2022 January to 2025 August.

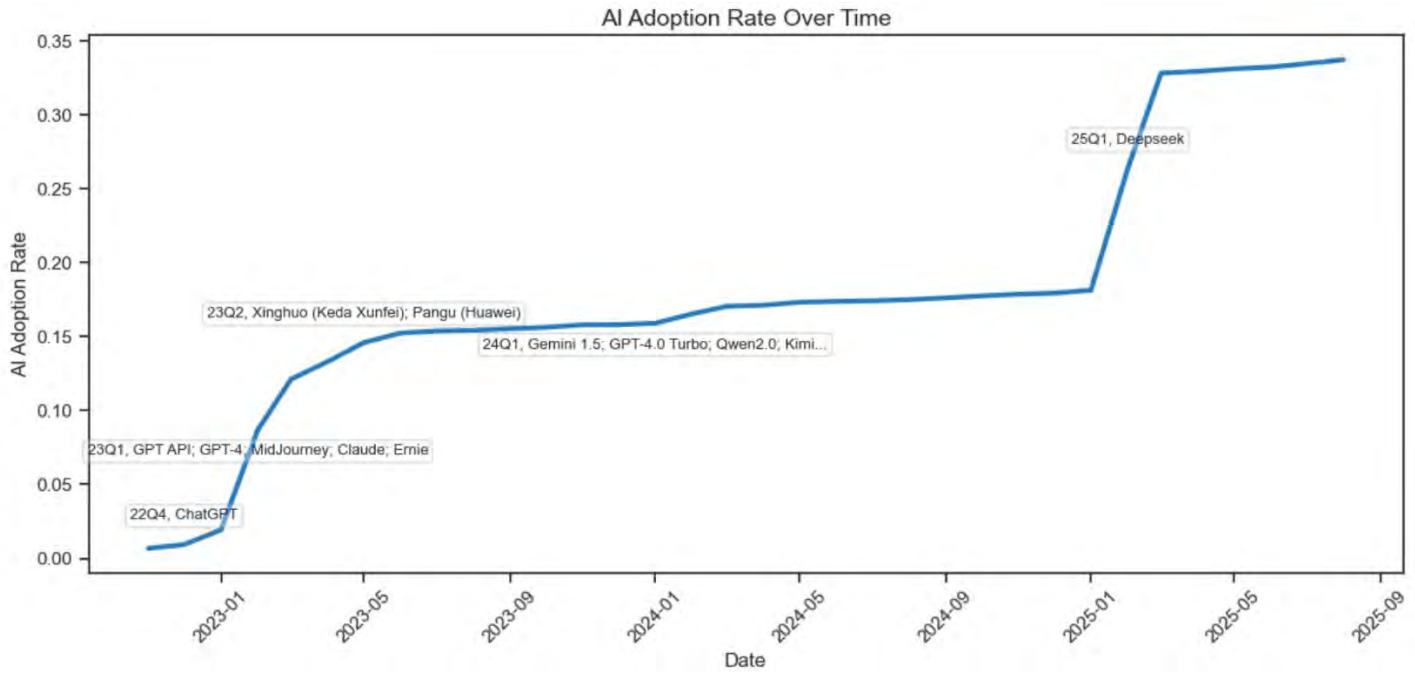


Figure 2. GenAI Adoption Over Time

This figure shows the time series of the GenAI adoption rate based on the answers on the investor interaction platforms. The sample period is from 2022 January to 2025 August. Firms that are questioned in the IIPs are included in the sample, which include 3656 firms. AI Adoption Rate is the percentage of firms that has responded GenAI adoption by the time divided by the total number of firm

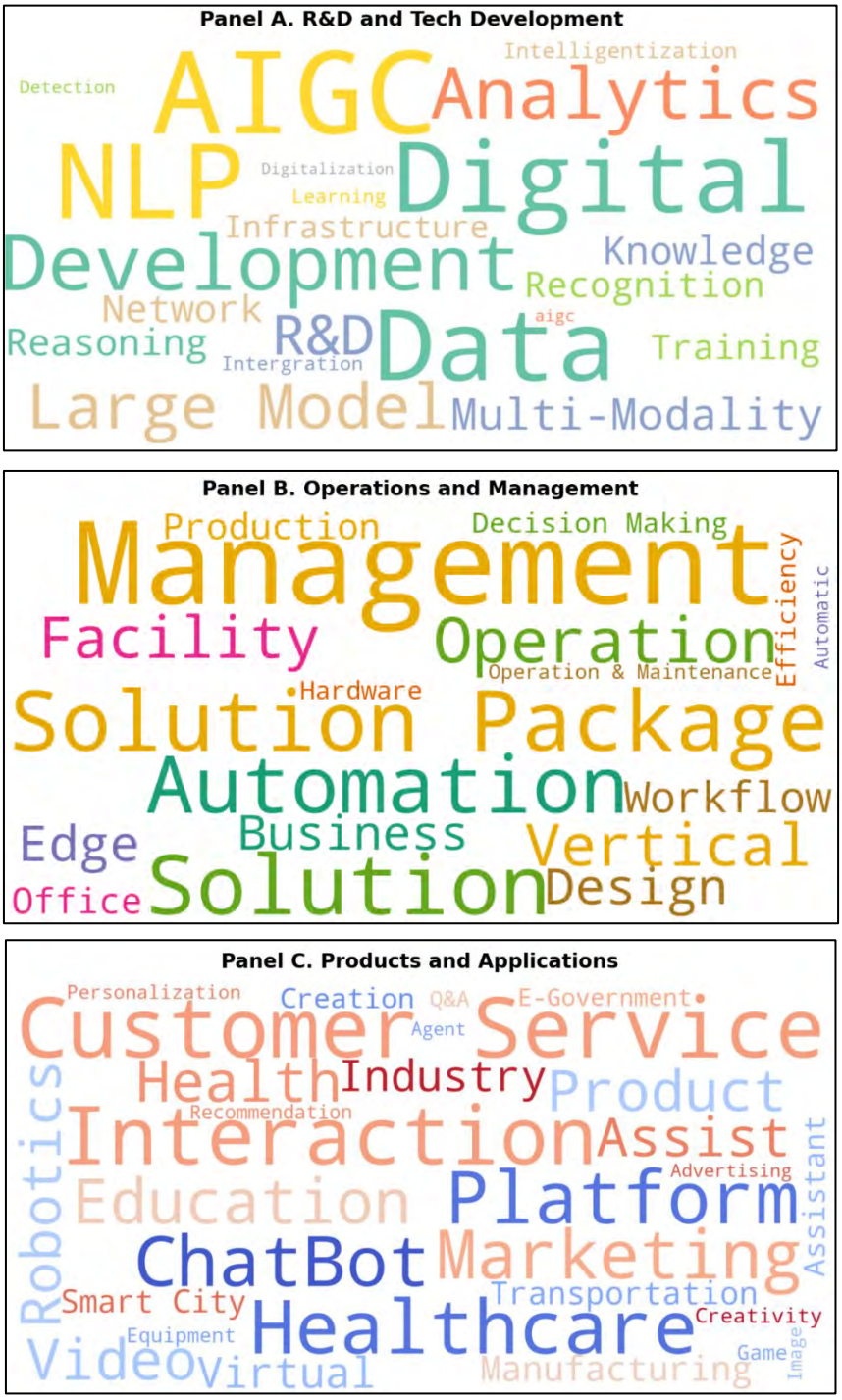


Figure 3. Word Clouds of GenAI Applications

This figure presents word clouds illustrating the most frequently mentioned GenAI application terms in the IIP responses. We first use DeepSeek V3 to summarize each firm's reported applications, then apply LDA to extract 80 representative and distinctive keywords. These keywords are classified into three categories: R&D and Technology Development, Operations and Management, and Products and Applications. Panel A displays the word cloud for R&D and Technology Development, Panel B for Operations and Management, and Panel C for Products and Applications. The size of each word reflects its frequency in the summarized application descriptions.

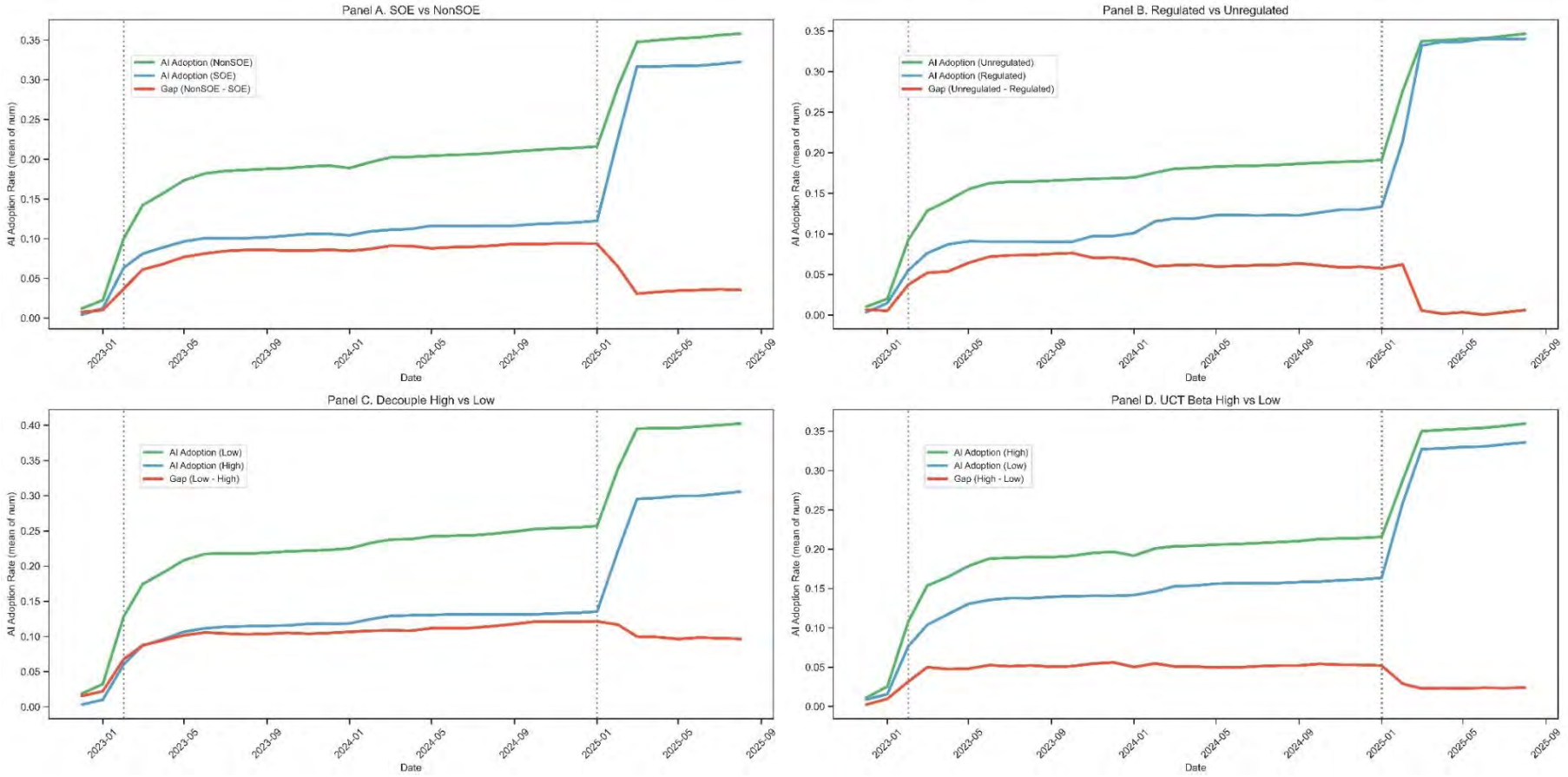


Figure 4. Adoption Delay of Sovereign and Geopolitically Important Firms

This figure plots the AI adoption rates of sovereign and non-sovereign firms from December 2022 to August 2025, based on firm responses from the investor interaction platform. Four sovereignty proxies are used, including state-owned enterprises (SOEs), Regulated Industries, Decoupling and stock return's US-China Tension Beta. Each panel plots the AI Adoption of sovereign and non-sovereign firms and the gap between two groups. The first dashed line is at February 2023, and the second dashed line is at February 2025.

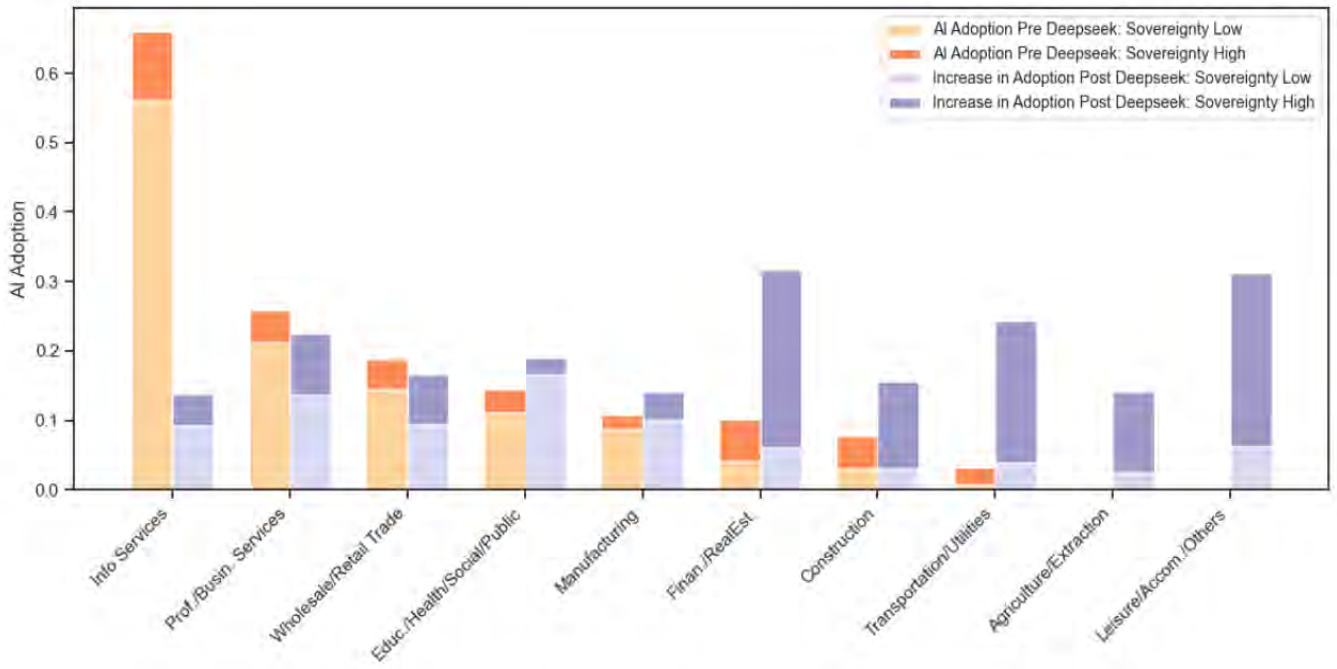


Figure 5. Heterogeneity of Adoption Delay at Industry Grouping Level

This figure shows the AI adoption rate before DeepSeek (orange bars) and the increase in adoption post-DeepSeek (purple bars) across industry groups, sorted by the pre-DeepSeek adoption rate. The pre-DeepSeek adoption rate is the share of firms that reported adopting AI before January 2025 among those that were asked about their AI adoption. The increase in adoption post-DeepSeek is the gap between AI adoption rate as of August 2025 minus the adoption rate as of January 2025. Industry groupings follow the classification methodology of Bick et al. (2024) to ensure comparability. The industry definitions are as follows: Info Service includes CSC industries Information Technology and Media; Prof./Busin. Services includes Business Services and Professional Services; Finan./RealEst. includes Finance and Real Estate; Leisure/Accomo./Others include Leisure, accommodation and others. Educ./Health/Social/Public include Education, Health and Social/Public Services. Transportation/Utilities include Transportation and Utilities; Construction refers to CSC industry E; Manufacturing refers to CSC industry C; Agriculture/Extraction includes Agriculture and Mining/Extraction.

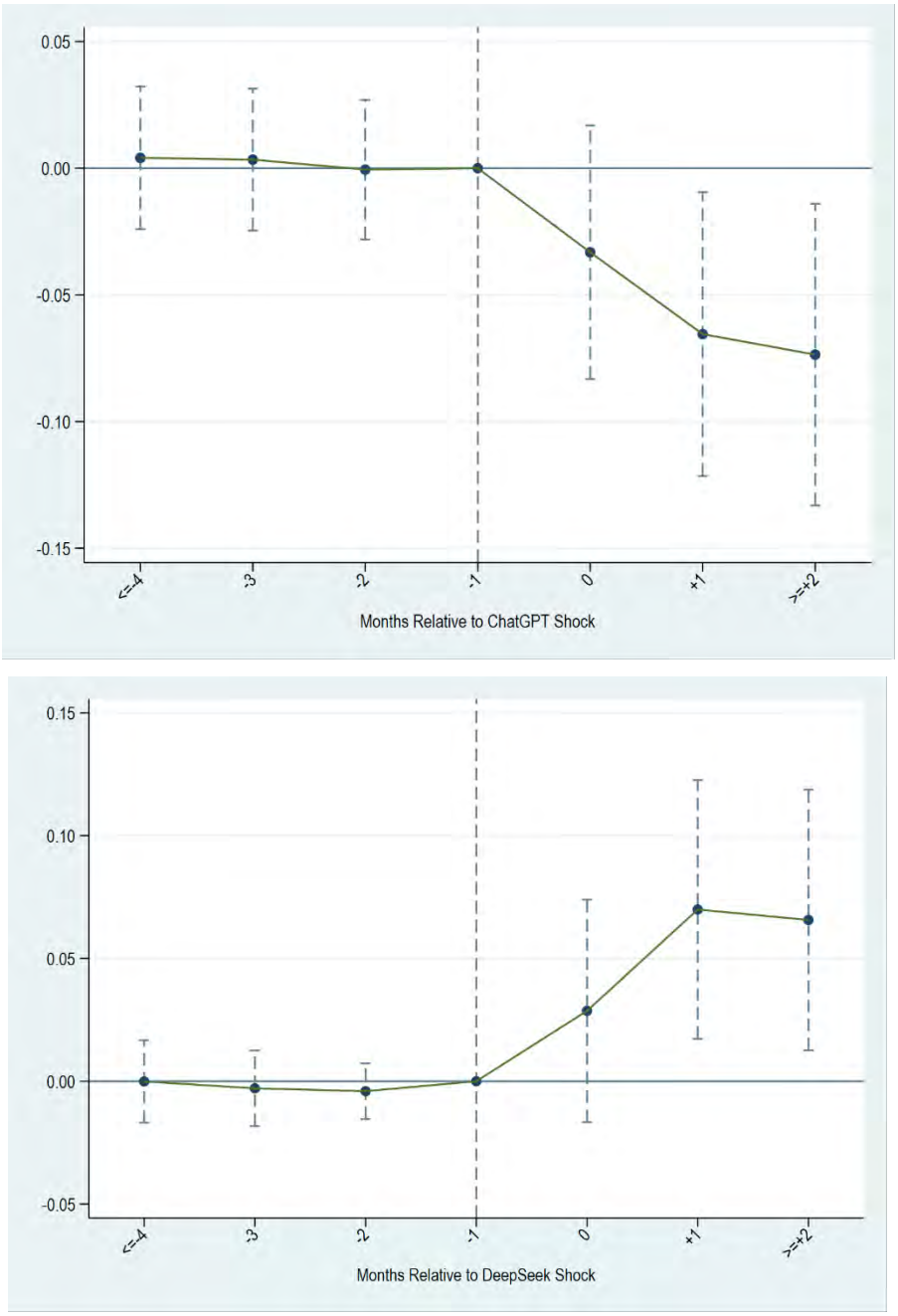


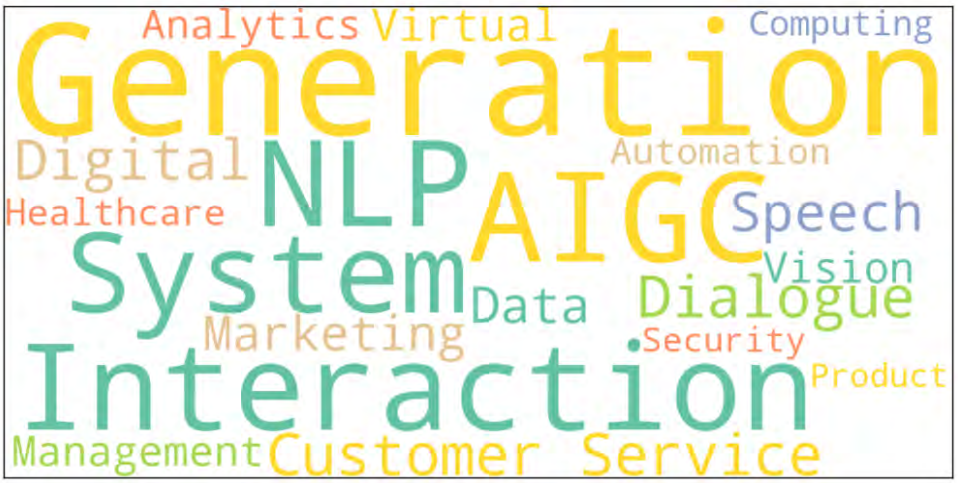
Figure 6. Dynamic Effects of Model Release on Firms' GenAI Adoption

This figure presents dynamic treatment effects of ChatGPT and DeepSeek releases on firms' GenAI adoption, respectively. We include a series of interaction terms, $HighSovIndex \times Post_k$, in the DID regression. For the ChatGPT analysis, 2023 January, one month prior to the ChatGPT shock, serves as the benchmark. For the DeepSeek analysis, 2025 January, one month prior to the DeepSeek shock, serves as the benchmark. We plot the coefficients on the interaction terms and their 95% confidence intervals. All regressions include firm fixed effects, time-industry fixed effects and time-city fixed effects. Standard errors are clustered at the firm level.

Figure 7. Word Clouds of Application Keywords in Answers by Legal Risks

This figure presents word clouds of the most frequently occurring terms derived from IIPs answers associated with high and low data breach legal risks, respectively. The size of each word reflects its relative frequency in the AI application summaries extracted from the IIP responses. Specifically, answers confirming AI adoption are processed by DeepSeek V3 to generate concise summaries of the mentioned applications. Based on these summaries, we apply Latent Dirichlet Allocation (LDA) to identify the 20 most representative and distinctive keywords. Panel A displays the word cloud for answers classified as high-risk, while Panel B corresponds to those with low-risk. The answer-level data breach legal risks scores are generated using LLM based on a structured evaluation prompt.

Panel A. GPT



Panel B. DeepSeek

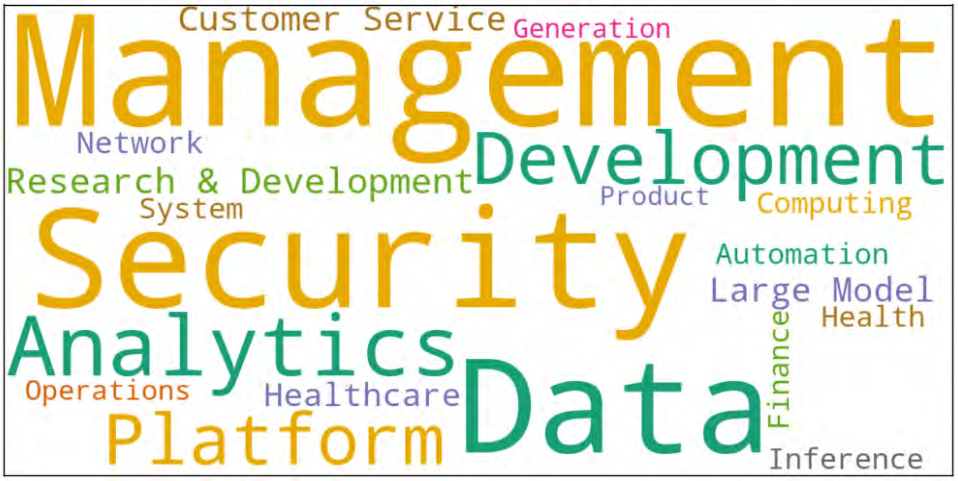


Figure 8. Word Clouds of Application Keywords in Answers by Model

This figure presents the word clouds of the most frequently used words based on IIP answers with GPT adoption and DeepSeek adoption, respectively. The size of each word in the cloud is proportional to the frequency of the word in the application summaries of IIPs answers. The answers in IIPs are input to DeepSeek V3 to generate summary of the mentioned applications. Then based on the summary, we use LDA to extract the most representative and distinctive 20 keywords. The word cloud in Panel A is based on the answers with GPT adoption and the word cloud in Panel B is based on the answers with DeepSeek adoption.

Table 1. Summary Statistics

This table reports the summary statistics. Panel A presents descriptive statistics for the full firm-month sample from January 2022 to August 2025, covering 3,656 firms that received at least one GenAI-related question on the investor interaction platform. Panel B reports the correlation matrix for the four sovereignty proxies (*** denotes significance at the 1% level). Panel C summarizes sovereignty and financial characteristics for the ChatGPT and DeepSeek samples. All variables are defined in Appendix A.

Panel A. Summary Statistics of Main Variables

Variable	N	Mean	SD	p50
SOE	151285	0.292	0.454	0.000
RegInd	153598	0.081	0.273	0.000
Decouple	132696	0.900	0.029	0.894
UCTBeta	131006	-0.005	0.047	-0.003
AI_Adoption	153598	0.144	0.351	0.000
Adopt_GPT	153598	0.074	0.261	0.000
Adopt_DeepSeek	153598	0.037	0.188	0.000
Log(1+#AI)	151285	10.800	21.700	2.000
I(AIJobPosting>0)	153598	0.159	0.366	0.000
AIJobPosting	153598	0.479	1.591	0.000
GovtSpendtoGDP	149386	0.193	0.080	0.189
IfBeijing	153598	0.099	0.299	0.000
SovIndex	149389	0.153	0.360	0.000
HHI	153596	0.053	0.057	0.035
Log(1+#DataAssets)	151252	3.584	0.754	3.434
DataBreachRisk	83422	3.327	0.812	4.000
SA	147625	-3.951	0.315	-3.969
WW	147589	-1.034	0.377	-1.019
KZ	134898	1.231	2.634	1.399

Panel B. Correlation Matrix of the Four Sovereignty Proxies

	SOE	*RegInd*	*HighDecouple*	*HighUCTBeta*
SOE	1			
RegInd	0.270***	1		
HighDecouple	0.064***	0.067***	1	
HighUCTBeta	0.064***	0.016***	0.047***	1

Panel C. Characteristics of Firms in the GPT Sample and DeepSeek Sample

	GPT Sample	DeepSeek Sample
SOE	23.44%	26.28%
RegInd	5.74%	5.96%
HighDecouple	50.07%	48.53%
HighUCTBeta	56.19%	51.70%
HighSovIndex	27.50%	28.84%
Log(AT)	22.31	22.43
MB	3.91	3.31
Age	21.89	23.57

Table 2. Validating AI Adoption Measures.

This table presents the results validating our GenAI adoption measures. Panel A presents the relationship between our AI adoption measure and firms' GenAI-related behaviors. The dependent variables are defined as follows: *Log(1+#AI)* is the natural logarithm of one plus the number of AI-related keywords mentioned in the MD&A section of the annual report; *I(Algorithm>0)* is a dummy variable equal to 1 if the firm filed at least one internet information service algorithm with regulators in that year, and 0 otherwise; *I(AIJobPosting>0)* is a dummy variable equal to 1 if the firm posted at least one AI-related job position, and 0 otherwise. The main independent variable, *AI_Adoption*, equals 1 if the firm's response in any dialogue during or before month *t* indicates GenAI adoption, and 0 otherwise. The sample is a firm-year panel from 2022 to 2024. All regressions include year fixed effects, and the even-numbered columns additionally control for the lagged dependent variable. The *t*-statistics, shown in parentheses, are calculated using standard errors clustered by firm. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. Panel B reports the agreement between LLM and human expert annotations for a random validation sample of 50 Q&A items. The confusion-matrix layout presents true positives (TP), false negatives (FN), false positives (FP) and true negatives (TN). Recall = TP / (TP + FN), while precision = TP / (TP + FP). Overall accuracy and the F1 score are reported in the bottom row.

Panel A. GenAI Adoption Measure and Firm's GenAI-related behaviors.

	(1)	(2)	(3)	(4)	(5)	(6)
	Log(1+#AI)		*I(Algorithm>0)*		*I(AIJobPosting>0)*	
AI_Adoption	1.654***	0.169***	0.032***	0.029***	0.178***	0.094***
	(29.76)	(8.79)	(6.08)	(6.12)	(10.88)	(7.51)
$Log(1+\#AI)_{t-1}$		0.888***				
		(192.41)				
$I(Algorithm>0)_{t-1}$				0.633***		
				(3.30)		
$I(AIJobPosting>0)_{t-1}$						0.440***
						(43.26)
Year FE	YES	YES	YES	YES	YES	YES
Observations	10,633	10,222	10,977	10,977	10,977	10,977
R-squared	0.149	0.811	0.028	0.076	0.014	0.190

Panel B. Agreement between LLM and Human Annotation.

	LLM Predicts Adoption	LLM Predicts Non-Adoption	Metric
Human Predicts Adoption	20	2	Recall = TP / (TP + FN) = 0.91
Human Predicts Non-Adoption	2	26	Precision = TP / (TP + FP) = 0.91
Overall			Accuracy = 0.92 F1 Score = 0.91

Table 3. Sovereignty and GenAI Adoption: Individual Sovereignty Proxies

This table examines the effects of sovereignty on firms' GenAI adoption using four individual sovereignty proxies. The dependent variable, *AI_Adoption*, equals 1 if the firm's response in any dialogue during or before month *t* indicates GenAI adoption, and 0 otherwise. *SovProxy* represents one of four sovereignty proxies—*SOE*, *RegInd*, *HighDecouple*, or *HighUCTBeta*. *SOE* equals 1 if the firm is ultimately controlled by a state entity and 0 otherwise. *RegInd* equals 1 if the firm operates in a sector the state designates as strategically important, and 0 otherwise. *HighDecouple* equals one if the firm's decoupling index is above the sample median and zero otherwise. We follow Han et al. (2024) and use patent citation data to construct the decoupling index that measures how frequently Chinese patents cite U.S. versus domestic patents. *HighUCTBeta* equals one if the firm's U.S.-China tension beta (*UCTBeta*) is above the sample median and zero otherwise. *UCTBeta* is estimated from regressing monthly stock returns on changes in the U.S.-China tension index and the Fama-French three factors. Panel A reports the effects of the ChatGPT release using data from August 2022 to August 2023. *PostGPT* equals 1 for months on or after February 2023, when the ChatGPT Plus was released for stable business and retail use and gained widespread attention. Panel B reports the effects of the DeepSeek release using data from August 2024 to August 2025. *PostDS* equals 1 for months on or after Feburary 2025, when DeepSeek R1 was released and gained widespread attention. Panel C presents the results of horse race regressions including all sovereignty proxies. All specifications include firm fixed effects and year-month fixed effects. The *t*-statistics, shown in parentheses, are calculated using standard errors clustered by firm. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A. GPT Shock

	AI_Adoption			
	(1)	(2)	(3)	(4)
SovProxy=	*SOE*	*RegInd*	*HighDecouple*	*HighUCTBeta*
SovProxy×PostGPT	-0.058**	-0.025	-0.089***	-0.056***
	(-2.43)	(-0.57)	(-4.41)	(-2.72)
Firm FE	YES	YES	YES	YES
Time FE	YES	YES	YES	YES
Observations	19,012	21,037	19,474	19,474
R-squared	0.627	0.595	0.622	0.620

Panel B. DeepSeek Shock

	AI_Adoption			
	(1)	(2)	(3)	(4)
SovProxy =	*SOE*	*RegInd*	*HighDecouple*	*HighUCTBeta*
SovProxy×PostDS	0.086***	0.076*	0.032*	0.044**
	(3.86)	(1.76)	(1.78)	(2.49)
Firm FE	YES	YES	YES	YES
Time FE	YES	YES	YES	YES
Observations	19,379	19,542	19,542	19,542
R-squared	0.858	0.856	0.856	0.856

Panel C. Horse Race of the Four Sovereignty Proxies

	(1)		(2)
	AI_Adoption		*AI_Adoption*
SOE×PostGPT	-0.051**	*SOE×PostDS*	0.077***
	(-2.11)		(3.38)
RegInd×PostGPT	0.003	*RegInd×PostDS*	0.045
	(0.07)		(1.02)
HighUCTBeta×PostGPT	-0.079***	*HighUCTBeta×PostDS*	0.022
	(-3.83)		(1.24)
HighDecouple×PostGPT	-0.050**	*HighDecouple×PostDS*	0.039**
	(-2.41)		(2.24)
Firm FE	YES	Firm FE	YES
Time FE	YES	Time FE	YES
Observations	19,012	Observations	19,379
R-squared	0.631	R-squared	0.859

Table 4. Sovereignty and GenAI Adoption: PCA-Based Sovereignty Index

This table examines the effects of sovereignty on firms' GenAI adoption using a unified PCA-based sovereignty index. The dependent variable, *AI_Adoption*, equals 1 if the firm's response in any dialogue on IIPs during or before month *t* indicates GenAI adoption, and 0 otherwise. *HighSovIndex* is a dummy variable that equals one if the PCA-based sovereignty index is above the sample median and zero otherwise. This PCA-based index is constructed from the four individual sovereignty proxies using principal component analysis (PCA). Columns (1) and (2) report the effects of the ChatGPT release, and columns (3) and (4) report the effects of the DeepSeek release. Other variables follow Table 3. All regressions include firm and time fixed effects; Columns (2) and (4) additionally include time–industry and time–city fixed effects. The *t*-statistics, shown in parentheses, are calculated using standard errors clustered by firm. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)		(3)	(4)
	AI_Adoption			*AI_Adoption*	
HighSovIndex×PostGPT	-0.052**	-0.062**	*HighSovIndex×PostDS*	0.085***	0.063**
	(-2.29)	(-2.19)		(3.97)	(2.38)
Firm FE	YES	YES	Firm FE	YES	YES
Time FE	YES	YES	Time FE	YES	YES
Time-Industry FE	NO	YES	Time-Industry FE	NO	YES
Time-City FE	NO	YES	Time-City FE	NO	YES
Observations	19,542	19,366	Observations	21,037	19,463
R-squared	0.857	0.872	R-squared	0.596	0.685

Table 5. Alternative Explanation: Preference for Open-Source Model

This table presents the effects of sovereignty on firms' adoption of open-source versus closed-source GenAI models. We focus on the releases of Gemini (closed-source) and LLaMA (open-source). The sample covers a 13-month window (−6m, +6 m) around each model's release. For Gemini, the event date is December 2023, when Google released Gemini 1.0; for LLaMA, the event date is July 2023, when Meta released LLaMA2. The dependent variables are specific model adoption, defined as follows: Gemini is a dummy variable equals to 1 if the firm has indicated Gemini adoption in its platform responses by time *t*. *Adopt_LLaMA* is a dummy variable equals to 1 if the firm has indicated LLaMA adoption in its platform responses by time *t*. *HighSovIndex* is a dummy variable that equals one if the PCA-based sovereignty index is above the sample median and zero otherwise. *PostGemini* and *PostLLaMA* are indicators for the post-release periods of Gemini and LLaMA, respectively. All specifications include firm, time-industry and time-city fixed effects. The *t*-statistics, shown in parentheses, are calculated using standard errors clustered by firm. ∗∗∗, ∗∗, and ∗ indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Adoption_Gemini	*Adoption_LLaMA*
	(1)	(2)
HighSovIndex×PostGemini	-0.002*	
	(-1.95)	
HighSovIndex×PostLLaMA		-0.002**
		(-2.19)
Firm FE	YES	YES
Time-Industry FE	YES	YES
Time-City FE	YES	YES
Observations	40,667	40,121
R-squared	0.497	0.607

Table 6 Alternative Explanation: Preference for Low-Cost Model

This table examines how model costs and financial constraints shape firms' GenAI adoption. The dependent variable, *AI_Adoption*, equals 1 if any dialogue on the IIPs during or before month t indicates GenAI adoption, and 0 otherwise. *HighSovIndex* equals 1 if the PCA-based sovereignty index is above the sample median. *HighChar* captures high cost sensitivity in column (1) and high financial constraints in columns (2)–(4). *HighCostBeta* equals 1 if the firm's cost sensitivity—measured as the percentage difference between output prices and marginal cost—is below the sample median. Financial-constraint proxies include SA (Hadlock and Pierce, 2010), KZ (Kaplan and Zingales, 1997), and WW (Whited and Wu, 2006). *HighSA*, *HighKZ*, and *HighWW* equal 1 if the respective index is above the sample median. Panel A reports results for the ChatGPT release, and Panel B reports results for the DeepSeek release. Other variables are defined as in Table 3. All regressions include firm, time–industry and time–city fixed effects. The *t*-statistics (in parentheses) are based on standard errors clustered at the firm level.

Panel A. GPT Shock

	AI Adoption			
	(1)	(2)	(3)	(4)
HighChar=	*HighCostBeta*	*HighSA*	*HighKZ*	*HighWW*
HighSovIndex×PostGPT	-0.064**	-0.064**	-0.067**	-0.066**
	(-2.17)	(-2.20)	(-2.17)	(-2.29)
HighChar×PostGPT	-0.003	-0.016	-0.023	-0.025
	(-0.14)	(-0.71)	(-0.96)	(-1.15)
Firm FE	YES	YES	YES	YES
Time-Industry FE	YES	YES	YES	YES
Time-City FE	YES	YES	YES	YES
Observations	18,288	18,783	16,655	18,783
R-squared	0.689	0.689	0.698	0.690

Panel B. DeepSeek Shock

	AI Adoption			
	(1)	(2)	(3)	(4)
HighChar=	*HighCostBeta*	*HighSA*	*HighKZ*	*HighWW*
HighSovIndex×PostDS	0.059**	0.061**	0.051*	0.048*
	(2.28)	(2.26)	(1.91)	(1.82)
HighChar×PostDS	-0.015	0.028	-0.021	-0.042**
	(-0.76)	(1.41)	(-1.03)	(-2.10)
Firm FE	YES	YES	YES	YES
Time-Industry FE	YES	YES	YES	YES
Time-City FE	YES	YES	YES	YES
Observations	13,791	17,798	17,185	17,798
R-squared	0.884	0.873	0.876	0.873

Table 7 Cross-Sectional Tests: The Influence of State Power

This table reports how state power influences the baseline effects. The dependent variable, *AI_Adoption*, equals 1 if any dialogue on the IIPs during or before month *t* indicates GenAI adoption, and 0 otherwise. *HighSovIndex* equals 1 if the PCA-based sovereignty index is above the sample median. Columns (1) and (2) split the sample by the median level of local government spending (local fiscal expenditure over GDP). Columns (3) and (4) split firms by whether they are headquartered in Beijing. Columns (5) and (6) split firms by political connections, defined as having a general manager or board chair with prior employment in central or local government. Panel A reports results for the ChatGPT release, and Panel B reports results for the DeepSeek release. Other variables are defined as in Table 3. All regressions include firm, time-industry and time-city fixed effects The *t*-statistics (in parentheses) are calculated using standard errors clustered by firm. p-value for the difference in coefficients is obtained through bootstrap. ∗∗∗, ∗∗, and ∗ indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A. GPT Shock

	AI_Adoption					
	(1)	(2)	(3)	(4)	(5)	(6)
	Low GovtSpending	*High GovtSpending*	*NonBeijing*	*Beijing*	*PC=0*	*PC=1*
HighSovIndex×PostGPT	-0.046	-0.073*	-0.041	-0.169**	-0.047	-0.202**
	(-1.14)	(-1.66)	(-1.34)	(-2.00)	(-1.45)	(-2.24)
p-value Test Low=High	<0.10*		<0.01***		<0.01***	
Firm FE	YES	YES	YES	YES	YES	YES
Time-Industry FE	YES	YES	YES	YES	YES	YES
Time-City FE	YES	YES	YES	YES	YES	YES
Observations	10,962	8,382	16,921	2,399	15,985	2,710
R-squared	0.665	0.714	0.667	0.739	0.692	0.757

Panel B. DeepSeek Shock

	AI_Adoption					
	(1)	(2)	(3)	(4)	(5)	(6)
	Low GovtSpending	*High GovtSpending*	*NonBeijing*	*Beijing*	*PC=0*	*PC=1*
HighSovIndex×PostDS	0.036	0.111***	0.083**	0.115	0.033	0.147
	(0.95)	(2.86)	(2.23)	(1.55)	(1.14)	(1.57)
p-value Test Low=High	<0.01***		<0.05**		<0.01***	
Firm FE	YES	YES	YES	YES	YES	YES
Time-Industry FE	YES	YES	YES	YES	YES	YES
Time-City FE	YES	YES	YES	YES	YES	YES
Observations	9,975	8,099	15,617	2,431	14,515	2,467
R-squared	0.877	0.871	0.870	0.867	0.874	0.905

Table 8. Cross-Sectional Tests: The Influence of Market Competition

This table reports how market competition influences the baseline effects. The dependent variable, *AI_Adoption*, equals 1 if any dialogue on the IIPs during or before month *t* indicates GenAI adoption, and 0 otherwise. *HighSovIndex* equals 1 if the PCA-based sovereignty index is above the sample median. We partition the sample into terciles based on the Herfindahl–Hirschman Index (HHI). Firms in the top tercile constitute the high-HHI group, and those in the bottom tercile form the low group. Panel A reports results for the ChatGPT release, and Panel B reports results for the DeepSeek release. All regressions include firm, time-industry and time-city fixed effects. Other variables are defined as in Table 3. The *t*-statistics (in parentheses), are calculated using standard errors clustered by firm. p-value for the difference in coefficients is obtained through bootstrap.

Panel A: ChatGPT Shock

	AI_Adoption	
	(1)	(2)
	Low HHI	High HHI
HighSovIndex×PostGPT	-0.020	-0.076*
	(-0.38)	(-1.72)
p-value Test Low=High	<0.05**	
Firm FE	YES	YES
Time-Industry FE	YES	YES
Time-City FE	YES	YES
Observations	5,967	8,415
R-squared	0.642	0.717

Panel B: DeepSeek Shock

	AI Adoption	
	(1)	(2)
	Low HHI	High HHI
HighSovIndex×PostDS	0.032	0.099*
	(0.72)	(1.79)
p-value Test Low=High	<0.01***	
Firm FE	YES	YES
Time-Industry FE	YES	YES
Time-City FE	YES	YES
Observations	5,897	5,580
R-squared	0.877	0.873

Table 9. Data Security Risks: The Influence on the Adoption Decisions

This table reports how data security risks influence the baseline effects. The dependent variable, *AI_Adoption*, equals 1 if any dialogue on the IIPs during or before month *t* indicates GenAI adoption, and 0 otherwise. *HighSovIndex* equals 1 if the PCA-based sovereignty index is above the sample median. In Panel A, we partition the sample into terciles based on the frequency of data-related keywords in firms' MD&A sections. Firms in the top tercile constitute the high-data-asset group, and those in the bottom tercile form the low group. In Panel B, we partition the sample into terciles based on firms' exposure to data-breach legal risks, with the top tercile classified as the high-data-breach-risk group and the bottom tercile as the low group. To measure these risks, we use a structured prompt that instructs DeepSeek to analyze firms' reports and assess their data-breach legal exposure. Other variables are defined as in Table 3. Columns (1) and (2) report results for the ChatGPT release, and columns (3) and (4) report results for the DeepSeek release. All regressions include firm, time-industry and time-city fixed effects. The t-statistics (in parentheses) are calculated using standard errors clustered by firm. p-value for the difference in coefficients is obtained through bootstrap. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A. High Data Assets vs Low Data Assets

	AI_Adoption			*AI_Adoption*	
	(1)	(2)		(3)	(4)
	Low Data Assets	High Data Assets		Low Data Assets	High Data Assets
HighSovIndex×PostGPT	-0.015	-0.098**	*HighSovIndex×PostDS*	0.022	0.077*
	(-0.21)	(-2.03)		(0.30)	(1.92)
p-value Test Low=High	<0.01***		*p*-value Test Low=High	<0.05**	
Firm FE	YES	YES	Firm FE	YES	YES
Time-Industry FE	YES	YES	Time-Industry FE	YES	YES
Time-City FE	YES	YES	Time-City FE	YES	YES
Observations	9,363	9,599	Observations	8,578	8,898
R-squared	0.639	0.726	R-squared	0.855	0.858

Panel B. High Data Breach Legal Risk vs Low Data Breach Legal Risk

	AI_Adoption			*AI_Adoption*	
	(1)	(2)		(3)	(4)
	Low Data Risk	High Data Risk		Low Data Risk	High Data Risk
HighSovIndex×PostGPT	-0.048	-0.088**	*HighSovIndex×PostDS*	0.029	0.064*
	(-1.05)	(-2.22)		(0.84)	(1.95)
p-value Test Low=High	<0.05**		*p*-value Test Low=High	<0.05**	
Firm FE	YES	YES	Firm FE	YES	YES
Time-Industry FE	YES	YES	Time-Industry FE	YES	YES
Time-City FE	YES	YES	Time-City FE	YES	YES
Observations	8,400	10,285	Observations	8,547	10,624
R-squared	0.673	0.710	R-squared	0.921	0.899

Table 10. Data Security Risks: The Influence on the Deployment Choices

This table reports how data security risks influence adopting firms' deployment choices for GenAI models. Panel A presents a cross-sectional sample of firms that adopted either GPT or DeepSeek during the sample period. Using a prompt-based approach with DeepSeek V3, we evaluate firms' reported uses of GenAI models and assign each firm a data-security-risk score (ranging from 1 to 5). This score, DataRisk_Application, serves as the dependent variable. *HighSovIndex* equals 1 if the PCA-based sovereignty index is above the sample median. Columns (1) and (2) report results for firms that adopted GPT and DeepSeek, respectively. In Panel B, the sample is restricted to firms that adopted both GPT and DeepSeek. The independent variable, *AdoptDomAI*, equals 1 when the risk score corresponds to a DeepSeek application and 0 when it corresponds to a GPT application. The *t*-statistics (in parentheses) are calculated using standard errors clustered by industry. p-value for the difference in coefficients is obtained through bootstrap. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A. Cross-Firm Analysis

	DataRisk_Application	
	(1)	(2)
	Firms Adopting GPT	*Firms Adopting DeepSeek*
HighSovIndex	-0.301***	0.031
	(-4.27)	(0.33)
p-value (1) = (2)	<0.01***	
Observations	405	741
R-squared	0.018	0.000

Panel B. Within-Firm Analysis

	DataRisk_Application
HighSovIndex×AdoptDomAI	0.267**
	(2.48)
AdoptDomAI	-0.058
	(-1.46)
Firm FE	YES
Observations	396
R-squared	0.629

Table 11. AI Adoption and Market Reactions

This table examines the differential stock market reactions to firms' AI adoption announcements. The dependent variable is the cumulative abnormal return (CAR) around the announcement window[-1,+1], estimated under three asset pricing benchmarks: (i) CAPM-based CARs, (ii) Fama–French 3-factor model CARs, and (iii) Fama–French 5-factor model CAR. Panel A reports the average market reactions to all AI adoption announcements. *AI_Adoption* equals 1 if the firm announces GenAI adoption, and zero otherwise. Panel B explores cross-sectional differences by model type and sovereignty exposure. *AdoptDomAI* is a dummy variable equal to 1 if the firm adopts DeepSeek, and 0 if it adopts GPT. *HighSovIndex* equals 1 if the PCA-based sovereignty index is above the sample median. All specifications include year-month and weekday fixed effects. *t*-statistics (in parentheses) are based on standard errors clustered by announcement date. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Panel A. Market Reactions to GenAI Adoption

	(1)	(2)	(3)
	CAR_CAPM	*CAR_FF3*	*CAR_FF5*
AI_Adoption	0.007***	0.005***	0.005***
	(4.05)	(3.25)	(3.28)
Year-Month FE	YES	YES	YES
Weekday FE	YES	YES	YES
Observations	4,884	4,884	4,884
R-squared	0.044	0.029	0.029

Panel B. Sovereignty and Market Reactions to GenAI Adoption

	(1)	(2)	(3)
	CAR_CAPM	*CAR_FF3*	*CAR_FF5*
HighSovIndex×*AdoptDomAI*	0.012***	0.013***	0.012***
	(2.96)	(3.29)	(3.20)
AdoptDomAI	-0.018	-0.023	-0.022
	(-0.47)	(-0.68)	(-0.58)
HighSovIndex	-0.010***	-0.010***	-0.009***
	(-2.82)	(-2.81)	(-2.65)
Observations	685	685	685
R-squared	0.072	0.061	0.061
YearMonth FE	YES	YES	YES
Weekday FE	YES	YES	YES

Appendix A. Variable Definitions

Variable	Definitions
Sovereignty Proxies	
SOE	A dummy variable that equals 1 if the firm is controlled by the states and 0 otherwise.
RegInd	Equals one if the firm is in regulated industries for which the state strategically retains control, and zero otherwise. The regulated industries include mining, railroads, trucking, airlines, telecommunications, energy supply, banking and media. These industries are identified based on policy reports from the state council (The General Office of the State Council, 2006) following Bradshaw et al. (2019).
HighUCTBeta	Equals one if the firm's *UCT beta* is above the sample median and zero otherwise. *UCT beta* is the beta of stock returns on US-China Tension. The beta is estimated using monthly stock returns prior to 2023, requiring at least 24 observations. The UCT beta is obtained by regressing monthly stock returns on change over US-China Tension, controlling for Fama-French 5 factors, following Rogers et al., (2022).
HighDecouple	Equals one if the firm's *Decouple* is above the sample median and zero otherwise. *Decouple* is a measure of China-US decoupling from Han et al. (2024). It captures the firms' tendency to build upon U.S. patents and is constructed at the IPC (International Patent Classification) level. Each firm is assigned a decoupling value based on its most frequently filed IPC. Higher values of Decouple indicate a lower reliance on U.S. patent knowledge.
AI Adoption Measures	
AI Adoption	A dummy variable equal to 1 if the firm has indicated AI adoption in its platform responses by time t.
Adopt_GPT	A dummy variable equal to 1 if the firm has indicated the adoption of GPT-related models (e.g., ChatGPT, GPT-4) in its platform responses by time t.
Adopt_Sora	A dummy variable equal to 1 if the firm has indicated the adoption of Sora-related models in its platform responses by time t.
Adopt_Midjourney	A dummy variable equal to 1 if the firm has indicated the adoption of Midjourney in its platform responses by time t.
Adopt_Gemini	A dummy variable equal to 1 if the firm has indicated the adoption of Gemini-related models in its platform responses by time t.

Adopt_Claude	A dummy variable equal to 1 if the firm has indicated the adoption of Claude-related models in its platform responses by time t.
Adopt_DeepSeek	A dummy variable equal to 1 if the firm has indicated the adoption of DeepSeek in its platform responses by time t.
Adopt_Ernie	A dummy variable equal to 1 if the firm has indicated the adoption of Ernie (Wenxin Yiyan) in its platform responses by time t.
Adopt_Qwen	A dummy variable equal to 1 if the firm has indicated the adoption of Qwen-related models in its platform responses by time t.
Adopt_Doubao	A dummy variable equal to 1 if the firm has indicated the adoption of Doubao in its platform responses by time t.
Adopt_Kimi	A dummy variable equal to 1 if the firm has indicated the adoption of Kimi in its platform responses by time t.

Validating Variables

Log(1+#AI)	The natural logrithm of one plus the number of AI word counts in MD&A section of the firm's annual report.
I(AIJobPosting>0)	A dummy variable equals to 1 if the firm post at least one AI-related position at year month t. [12]

[12] The AI words are obtained in the following steps:

Initial Term Selection: A set of 52 core AI terms (such as "artificial intelligence," "machine learning," "Internet of Things," and "cloud computing") was manually compiled. This selection was informed by multiple sources, including the Chinese-translated list from Chen and Srinivasan (2020), industry analyses like Ping An Securities' *AI Industry Chain Overview*, reports from the China Business Industry Research Institute and Qianzhan Industry Research Institute, as well as the vocabulary published by the World Intellectual Property Organization (WIPO).

Semantic Expansion: Using the Word2vec model (Skip-gram architecture) as outlined in Mikolov et al. (2013) and applied similarly in Li et al. (2021), we trained on a corpus of annual reports and patent texts. For each seed word, we identified the 10 most semantically similar terms based on cosine similarity scores.

Refinement and Finalization: After eliminating duplicates, non-AI-related terms, and those with low frequency, we finalized a list of 73 terms that comprise the AI dictionary used in this study.

The AI words include (translated from Chinese): Artificial Intelligence, Computer Vision, Image Recognition, Knowledge Graph, Intelligent Education, Augmented Reality, Smart Governance, Feature Extraction, Business Intelligence, Smart Elderly Care, Support Vector Machine (SVM), Knowledge Representation, Pattern Recognition, Internet of Things (IoT), Human-Computer Dialogue, AI Products, Human-Computer Interaction, Data Mining, Smart Banking, Intelligent Customer Service, Virtual Reality, Autonomous Driving, Driverless Technology, Smart Finance, Big Data Marketing, Long Short-Term Memory (LSTM), AI Chips, Edge Computing, Cloud Computing, Deep Neural Networks, Deep Learning, Feature Recognition, Intelligent Insurance, Smart Retail, Intelligent Healthcare, Intelligent Transportation, Smart Home, Recurrent Neural Networks, Big Data Risk Control, Robotic Process Automation, Wearable Devices, Big Data Platforms, Augmented Intelligence, Big Data Operations, Machine Translation, Neural Networks, Speech Synthesis, Human-Machine Collaboration, Smart Agriculture, Smart Speakers, Convolutional Neural Networks, Question-Answering Systems, Reinforcement Learning, Big Data Analysis, Natural Language Processing (NLP), Big Data Management, Intelligent Computing, Voice Interaction, Machine Learning, Biometric Identification, Speech Recognition, Intelligent Regulation, Robo-Advisory, Intelligent Speech, Voiceprint Recognition, Facial Recognition, Intelligent Agents, Big Data Processing, Distributed Computing, Smart Sensors, Intelligent Search, Smart Environmental Protection.

I(Algo>0)	An indicator variable equal to one if a firm has filed at least one internet information service algorithm with regulators in a given year, and zero otherwise. Algorithm filings are mandated by two key regulations: the Provisions on the Administration of Algorithmic Recommendation for Internet Information Services (effective March 2022) and the Provisions on the Administration of Deep Synthesis of Internet Information Services (November 2022).

Cross Sectional Variables

GovtSpend	Government interventation is caluclated as the ratio of local fiscal spending to local GDP.
PC	*PC* is an indicator that equals one if either the general manager or the board chair was previously employed as a bureaucrat by the central government or a local government.
HHI	*Herfindahl–Hirschman Index* measures the degree of market concentration within an industry. For each industry-year, HHI is calculated as the sum of the squared market shares of all listed firms based on operating revenue. A higher HHI indicates a lower market competitiveness.
#Data in MD&A	The number of AI word counts in MD&A section of the firm's annual report. [13]

[13] Data words are obtained in a similar methodology to the AI words.

The data words include (translated from Chinese): Digital Infrastructure, Digital Facilities, Digital Factory, Digital Equipment, Digital Economy, Digital Technology Innovation, Digital Governance, Digital Management, Digital Operations & Maintenance (O&M), Digital Design, Digital Technological Means, Digital Technology, Digital Resources, Digital Factors, Digital Science and Technology, Key Digital Enterprises, Digital Applications, Data Infrastructure, Data Laboratory, Data Collection, Data Gathering, Data Development, Data Development and Utilization, Data Governance, Data Storage, Data-Driven, Data Processing, Data Computing, Data Utilization, Data Innovation, Data Standardization, Data Resource Management, Data Strategy, Data Resource Integration, Information Infrastructure, Information Facilities, Information Terminals, Information Equipment, Information Disclosure, Information Acquisition, Information Gathering, Information Management, Information Collection, Information Docking, Information R&D, Information Access, Information and Communications, Information Resources, Information Technology, Information Data, Information Science and Technology, Network Infrastructure, Network Information Systems, Network Data Centers, Network Facilities, Network Construction, Network Storage, Network Governance, Network Access, Network Data, Network Technology, Network Information Technology, Network Information, Network Digital, Network Services, Network Industry Chain, Network Performance, Network Establishment, Digital Ecosystem, Digital Capabilities, Digital Literacy, Digital Industry, Digital Space, Data Monitoring, Data Processing, Data Economy, Data Research, Data Protection, Data Interaction, Data Optimization, Data Aggregation, Data Talent Development, Data Encryption, Data Desensitization, Data R&D, Data Cleaning, Data Networking, Data Operations, Data Supply, Data Integration, Data Consolidation, Data Interfacing, Data Cloud Migration, Data Management, Data Access, Data Analysis, Data Centralized Management, Data Resources, Data Factors, Data Assets, Data Technology, Data Talent, Key Data Technologies, Data Experts, Data Algorithms, Data-Driven Intelligence, Data Information Technology, Data Services, Data Networks, Data Talent Teams, Data Computing Power, Data Engineering, Information Networks, Information Carriers, Information Operations, Information Manufacturing, Information Protection, Information Inclusion, Information Identification, Information Monitoring, Information Recording, Network Optimization, Network Functions, Network Applications, Network Expansion, Network Layout, Network Operations, Network Operation, Data Engineering Technology, Data Enterprises, Data SMEs (Small and Medium Enterprises), Leading Data Enterprises, Data Industry Bases, Data Ecosystem, Data Visualization, Data Resource Advantage, Data Space, Data Audit, Data Query, Data Research Institutes, Data Industrial Parks, Data Verification, Data Industry Layout, Data Scale, Data Capabilities, Data Utilization, Data Simulation, Data Interface, Data Security Levels, Data Acquisition, Data Technology Industry, Data Confidentiality, Data Industry Chain, Data Projects, Data + Artificial Intelligence, Data Empowerment, Data Production, Data Usage, Data Industry, Data Applications, Data Development, Data Insurance, Digital Platform,

Data Breach Legal Risk	An index capturing firm-level exposure to data breach and compliance risks, derived from textual analysis of MD&A disclosures and evaluated with reference to major Chinese data governance laws and classification guidelines. [14]
SA	Hadlock and Pierece (2010) financial constraint index. SA Index for firm i in year t is computed as -0.737**Size*-0.043**Size*^2 - 0.04**Age*, where *Size* is log(inflation -adjusted book assets), and *Age* is the current year minus the first year that the firm is listed.
KZ	Kaplan and Zingales (1997) financial constraint index. KZ Index for firm i in year t is computed as -1.002**CF* + 3.139**TLTD*-39.368**TDIV* - 1.315**CASH* - 0.283**Q*, where *CF* is cash flow normaized by total asset, *TLTD* is the ratio of the long-term debt to total assets *CASH* is is cash normalized by totoal asset, *TDIV* is the ratio of total dividends to assets and *Q* is Tobin's *Q*.

Digital Trade, Digital Certification, Digital Consumption, Digital Products, Digital Currency, Digital Product Security, Data Platform, Data Trading Platform, Data Open Platform, Data Service Platform, Data Transactions, Data Openness, Data Circulation, Data Resource Sharing, Data Sharing, Data Cooperation, Data Collaboration, Data Delivery, Data Application Services, Data Hosting, Data Consumption, Data Property Rights, Data Usage Rights, Information Platform, Information Sharing, Information Services, Information Consumption, Information Resource Sharing, Information Interconnectivity, Information Service System, Information Exchange, Information Service Level, Information Service Platform, Online Transactions, Online Sales, Network Data Security, Network Information Security, Network Security, Network Risks, Network Provision, Network Services, Network Interoperability, Network Integration, Network Interconnection, Network Service Providers, Cyber Sovereignty, Data Security, Data and Information Security, Data Evaluation, Data Service Providers, Data Transparency, Data Regulation, Data Vendors, Data Interoperability, Data Exchanges, Data Market, Data Providers.

[14] The prompt is as follows (translated from Chinese):
Please act as an expert in data security, international compliance, and national security. Analyze the cross-border data leakage risk levels (on a scale from 1 to 5, with 5 indicating the highest risk) for the following industries, and briefly explain the reasoning behind your assessment. Present the results in a tabular format.
Please base your evaluation on the following definition:
Cross-border data breach risk refers to the possibility that, during international business operations or data transfers—through mechanisms such as data transmission, remote access, or overseas hosting—national critical information, trade secrets, or personal data may be exposed to foreign regulatory authorities, intelligence agencies, or commercial competitors.
Industries with higher risk levels typically exhibit the following characteristics:

- Involvement with national strategic resources, critical infrastructure, national defense, technological sovereignty, or essential public services and livelihood;
- Frequent cross-border data exchanges, international business cooperation, or technological dependencies (e.g., cross-border R&D, overseas cloud services, outsourcing);
- High geopolitical sensitivity or relevance to export control regulations;
- Data leaks may trigger regulatory accountability, diplomatic disputes, supply chain disruptions, or significant reputational and economic losses.

Please assess the following industries:
\n\n{industry name}\n\n
Your output must follow this format (do not include the text in parentheses):
(Risk Level) 1/2/3/4/5 + (Justification) A brief explanation
For example:
4 + This industry involves critical infrastructure, has frequent cross-border data flows, and is significantly influenced by geopolitical factors.

WW	Whited and Wu (2006) financial constraint index. WW Index for firm i in fiscal year t is computed as -0.091**CF* - 0.062**Dividend* +0.021*Leverage - 0.044**Size*+0.102**IndustrySalesGrowth* - 0.035**FirmSalesGrowth*, where *CF* is operating cash flows normalized by total asset, dividend is an indicator that equals one if the firm pays cash dividends, leverage is the ratio of long-term debt to total assets, size is log of total assets, and industry sales growth is the average sales growth of all firms in the same CSC industry.

Appendix B. Prompts for Developing the AI Adoption Measure

You are a professional investor with financial literacy and a deep understanding of artificial intelligence (AI) and large model applications. Please judge whether the following content indicates that the company has applied artificial intelligence, and extract the specific AI names.

If a company plans or has already fine-tuned, trained, or integrated GenAI into its business, then you should consider the company as deploying GenAI; If a company beats around the bush and does not directly respond to its usage of GPT and other GenAI models, then you should consider it as not deploying GenAI;

If the company explicitly mentions usage of GenAI models such as GPT, DeepSeek, Qwen, please output the specific AI names; possible options include (but are not limited to): 'GPT', 'Sora', 'llama', 'claude', 'gemini', 'deepmind', 'grok','Dalle', 'Midjourney', 'Stable Diffusion' and 'Deepseek', etc; If the company vaguely mentions AI-related technologies but does not specify the name, then output "AI"; If the company explicitly states that it has not used AI, or only uses traditional machine learning technologies, please output " none ".

Please analyze the following content:

\n\n{reply_content}\n\n

Your output must be in the following format (do not output the content in parentheses):

（about AI Adoption）Deploy/NotDeploy+（The specific model）GPT, DeepSeek, Qwen, AI,None

For example: Deploy+AI; Deploy+Deepseek; NotDeploy+None;

Appendix C. Additional Tables

Table A1. Top 5 U.S. and Chinese AI Models Adopted by the Firms
This table presents the five US and China most frequently adopted AI models, based on responses from the investor interaction platform. Models are grouped by their country of origin (United States or China) and ranked by the number of firms explicitly claiming adoption of the specific model. Panel A lists U.S.-based models, while panel B lists models originating from China. *#Adopted* is the number of firms explicitly claiming adoption of the specific model. *#Questioned* is the number of questions containing keywords related to the specific model.

Panel A. Top U.S. Models

Country	Model	#Adopted	#Questioned
US	*GPT*	410	1541
US	*Sora*	35	412
US	*Llama*	21	41
US	*MidJourney*	20	34
US	*Gemini*	6	48

Panel B. Top Chinese Models

Country	Model	#Adopted	#Questioned
CN	*DeepSeek*	855	2702
CN	*Baidu*	152	396
CN	*Doubao*	53	379
CN	*Pangu*	46	385
CN	*ChatGLM*	43	54

Table A2. Characteristics of Firms Receiving GenAI-Related Questions vs. Firms Not Receiving Them

This table compares firm-level characteristics between firms that were asked GenAI-related questions (Asked = 1) and those not asked (Asked = 0). *AIwordsMD&A* measures the average number of AI-related keywords mentioned in annual report MD&A sections. *IfAlgorithmFilings* is an indicator equal to one if the firm has ever submitted algorithm-related filings, and zero otherwise; *NumAlgorithmFilings* is the total number of such filings. *IfAIJobPostings* equals one if the firm has ever posted AI-related job vacancies; *NumAIJobPostings* represents the total number of such postings. *t*-statistics for the difference in means are reported in the third column, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively.

	Asked=1	Asked=0	t-stats of Diff.
AIwordsMD&A	11.749	2.825	12.74***
IfAlgorithmFilings	0.014	0.000	4.88***
NumAlgorithmFilings	0.474	0.000	2.65***
IfAIJobPostings	0.773	0.630	11.10***
NumAIJobPostings	26.107	11.508	6.38***

Table A3. Robustness: Alternative AI Adoption Measures

This table reports robustness tests that use the adoption of specific AI models as the dependent variable. In column (1), *Adopt_GPT* equals 1 if a firm has reported GPT adoption in the IIPs by time *t*. In column (2), *Adopt_DeepSeek* equals 1 if a firm has reported DeepSeek adoption in the IIPs by time *t*. *HighSovIndex* is a dummy variable that equals one if the PCA-based sovereignty index is above the sample median and zero otherwise. Column (1) reports results for the ChatGPT shock, covering the period from August 2022 to August 2023. *PostGPT* equals 1 for months on or after February 2023. Column (2) reports results for the DeepSeek shock, covering the period from August 2024 to August 2025. *PostDS* equals 1 for months on or after February 2025. All regressions include firm, time–industry and time–city fixed effects. The *t*-statistics, shown in parentheses, are calculated using standard errors clustered by firm. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) *Adopt_GPT*		(2) *Adopt_DeepSeek*
HighSovIndex×PostGPT	-0.050**	*HighSovIndex×PostDS*	0.062*
	(-2.00)		(1.67)
Firm FE	YES	Firm FE	YES
Time-Industry FE	YES	Time-Industry FE	YES
Time-City FE	YES	YearMonth FE	YES
Observations	19,366	Observations	11,174
R-squared	0.634	R-squared	0.644

Table A4. Robustness: Alternative Samples

This table reports robustness tests using alternative samples. The dependent variable, *AI_Adoption*, is a dummy variable equal to 1 if the firm has reported GenAI adoption in its IIP responses by time *t*. *HighSovIndex* is a dummy variable that equals one if the PCA-based sovereignty index is above the sample median and zero otherwise. In Alternative Sample 1, we exclude the firms that did not respond to any questions. In Alternative Sample 2, we further exclude cases where the responses contain no relevant information, defined as responses that merely direct investors to company announcements or public disclosures (e.g., "please refer to the company's announcement"). Alternative Sample 3 additionally reclassifies responses indicating a "plan to use" or "intention to adopt" AI as non-adoption, retaining only firms that explicitly state they "have adopted" as adopters. Panel A presents the results for the ChatGPT shock, covering the period from August 2022 to August 2023, where *PostGPT* equals 1 for months on or after February 2023. Panel B reports the results for the DeepSeek shock, covering the period from August 2024 to August 2025, where *PostDS* equals 1 for months on or after February 2025. All regressions include firm, time–industry, and time–city fixed effects. The *t*-statistics, shown in parentheses, are based on standard errors clustered by firm. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A. ChatGPT Shock

	(1)	(2)	(3)
		AI Adoption	
	Alternative Sample 1	Alternative Sample 2	Alternative Sample 3
HighSovIndex×PostGPT	-0.066**	-0.074***	-0.070**
	(-2.30)	(-2.68)	(-2.46)
Firm FE	YES	YES	YES
Time-Industry FE	YES	YES	YES
Time-City FE	YES	YES	YES
Observations	19,081	18,870	18,870
R-squared	0.686	0.678	0.683

Panel B. DeepSeek Shock

	(1)	(2)	(3)
		AI Adoption	
	Alternative Sample 1	Alternative Sample 2	Alternative Sample 3
HighSovIndex×PostDS	0.065**	0.063**	0.055**
	(2.46)	(2.38)	(2.12)
Firm FE	YES	YES	YES
Time-Industry FE	YES	YES	YES
Time-City FE	YES	YES	YES
Observations	18,078	18,078	18,078
R-squared	0.873	0.872	0.876

Table A5. Keywords List for Exposures to AI and Data Assets

This table reports the AI-related and data–asset-related keywords used to construct the corresponding measures. Panel A lists the AI-related keywords employed to identify AI-related job postings and to count AI-related terms in the MD&A sections of annual reports. Panel B lists the data–asset-related keywords used to measure firms' exposure to data security based on the frequency of such terms in the MD&A sections of annual reports.

Panel A. The List of AI-Related Keywords

English	Chinese	English	Chinese	English	Chinese
Artificial Intelligence	人工智能	Deep Neural Network (DNN)	深度神经网络	Voice Interaction	语音交互
Computer Vision	计算机视觉	AI Chip	AI 芯片	Machine Learning (ML)	机器学习
Image Recognition	图像识别	Deep Learning	深度学习	Biometrics	生物识别
Knowledge Graph	知识图谱	Feature Recognition	特征识别	Speech Recognition	语音识别
Intelligent Education	智能教育	Intelligent Insurance	智能保险	Intelligent Regulation	智能监管
Augmented Reality (AR)	增强现实	Smart Retail	智能零售	Robo-Advisor	智能投顾
Smart Governance / E-Government	智能政务	Intelligent Healthcare	智能医疗	Intelligent Speech	智能语音
Feature Extraction	特征提取	Smart Transportation	智能运输	Voiceprint Recognition	声纹识别
Business Intelligence (BI)	商业智能	Smart Home	智能家居	Face Recognition	人脸识别
Intelligent Elderly Care	智能养老	Recurrent Neural Network (RNN)	循环神经网络	Intelligent Agent	智能体
Support Vector Machine (SVM)	支持向量机（SVM）	Big Data Risk Control	大数据风控	Big Data Processing	大数据处理

Knowledge Representation	知识表示	Robotic Process Automation (RPA)	机器人流程自动化	Distributed Computing	分布式计算
Pattern Recognition	模式识别	Wearable Devices	可穿戴产品	Intelligent Sensors	智能传感器
Internet of Things (IoT)	物联网	Big Data Platform	大数据平台	Intelligent Search	智能搜索
Human–Computer Dialogue	人机对话	Augmented Intelligence	增强智能	Smart Environmental Protection	智能环保
AI Product	AI 产品	Big Data Operations	大数据运营	Long Short-Term Memory (LSTM)	长短期记忆（LSTM）
Human–Computer Interaction (HCI)	人机交互	Machine Translation	机器翻译	Intelligent Chip	智能芯片
Data Mining	数据挖掘	Neural Network	神经网络	Edge Computing	边缘计算
Smart Banking	智慧银行	Speech Synthesis	语音合成	Cloud Computing	云计算
Intelligent Customer Service	智能客服	Human–AI Collaboration	人机协同	Big Data Analytics	大数据分析
Virtual Reality (VR)	虚拟现实	Smart Agriculture	智能农业	Natural Language Processing (NLP)	自然语言处理
Autonomous Driving	自动驾驶	Smart Speaker	智能音箱	Big Data Management	大数据管理
Driverless	无人驾驶	Convolutional Neural Network (CNN)	卷积神经网络	Intelligent Computing	智能计算
Smart Finance	智慧金融	Question Answering System (QA System)	问答系统	Reinforcement Learning	强化学习
Big Data Marketing	大数据营销				

Panel B. The List of Data-Asset-Related Keywords

English	Chinese	English	Chinese	English	Chinese
Digital Infrastructure	数字基础设施	Network Industry Chain	网络产业链	Data SMEs	数据中小企业
Digital Facility	数字设施	Network Performance	网络性能	Data Leading Enterprises	数据龙头企业
Digital Factory	数字工厂	Network Establishment	网络建立	Data Industry Base	数据产业基地
Digital Equipment	数字设备	Digital Ecosystem	数字生态	Data Ecosystem	数据生态
Digital Economy	数字经济	Digital Capability	数字能力	Data Visualization	数据可视
Digital Technology Innovation	数字技术创新	Digital Literacy	数字素养	Data Resource Advantages	数据资源优势
Digital Governance	数字治理	Digital Industry	数字产业	Data Space	数据空间
Digital Management	数字管理	Digital Space	数字空间	Data Audit	数据审核
Digital Operation & Maintenance	数字运维	Data Monitoring	数据监测	Data Query	数据查询
Digital Design	数字设计	Data Processing	数据加工	Data Research Institute	数据研究院
Digital Technology Means	数字技术手段	Data Economy	数据经济	Data Industrial Park	数据产业园
Digital Technology	数字技术	Data Research	数据研究	Data Verification	数据核查
Digital Resources	数字资源	Data Protection	数据保护	Data Industrial Layout	数据产业布局
Digital Elements / Factors	数字要素	Data Interaction	数据交互	Data Scale	数据规模
Digital Science & Technology	数字科技	Data Optimization	数据优化	Data Capability	数据能力

Leading Digital Enterprises	数字骨干企业	Data Aggregation	数据归集	Data Application	数据运用
Digital Application	数字应用	Data Talent Training	数据人才培养	Data Simulation	数据模拟
Data Infrastructure	数据基础设施	Data Encryption	数据加密	Data Interface	数据接口
Data Laboratory	数据实验室	Data Desensitization	数据脱敏	Data Security Levels	数据安全等级
Data Acquisition	数据采集	Data R&D	数据研发	Data Acquisition	数据获取
Data Collection	数据收集	Data Cleaning	数据清洗	Data Technology Industry	数据技术产业
Data Development	数据开发	Data Networking	数据联网	Data Confidentiality	数据保密
Data Development & Utilization	数据开发利用	Data Operations	数据运营	Data Industry Chain	数据产业链
Data Governance	数据治理	Data Supply	数据供给	Data Projects	数据项目
Data Storage	数据存储	Data Integration	数据整合	Data Artificial Intelligence	数据人工智能
Data-Driven	数据驱动	Data Integration (System)	数据集成	Data Empowerment	数据赋能
Data Processing	数据处理	Data Docking	数据对接	Data Production	数据生产
Data Computing	数据计算	Data to Cloud	数据上云	Data Usage	数据使用
Data Utilization	数据利用	Data Management	数据管理	Data Industry	数据产业
Data Innovation	数据创新	Data Access	数据接入	Data Applications	数据应用
Data Standardization	数据标准化	Data Analysis	数据分析	Data Development	数据发展
Data Resource Management	数据资源管理	Data Centralized Management	数据集中管理	Data Insurance	数据保险
Data Strategy	数据战略	Data Resources	数据资源	Digital Platform	数字平台

Data Resource Integration	数据资源整合	Data Elements / Factors	数据要素	Digital Trade	数字贸易
Information Infrastructure	信息基础设施	Data Assets	数据资产	Digital Authentication	数字认证
Information Facility	信息设施	Data Technology	数据技术	Digital Consumption	数字消费
Information Terminal	信息终端	Data Talent	数据人才	Digital Product	数字产品
Information Equipment	信息设备	Key Data Technologies	数据关键技术	Digital Currency	数字货币
Information Disclosure	信息披露	Data Experts	数据专家	Digital Product Safety	数字产品安全
Information Acquisition	信息获取	Data Algorithms	数据算法	Data Platform	数据平台
Information Collection	信息收集	Data Information	数据信息	Data Trading Platform	数据交易平台
Information Management	信息管理	Data Information Technology	数据信息技术	Data Open Platform	数据开放平台
Information Gathering	信息采集	Data Business	数据业务	Data Service Platform	数据服务平台
Information Docking	信息对接	Data Network	数据网络	Data Trading	数据交易
Information R&D	信息研发	Data Talent Team	数据人才队伍	Data Opening	数据开放
Information Access	信息接入	Data Computing Power	数据算力	Data Circulation	数据流通
Information & Communication	信息通信	Data Engineering	数据工程	Data Resource Sharing	数据资源共享
Information Resources	信息资源	Information Network	信息网络	Data Sharing	数据共享
Information Technology	信息技术	Information Carrier	信息载体	Data Cooperation	数据合作
Information Data	信息数据	Information Operation	信息运行	Data Collaboration	数据协同
Information Science & Technology	信息科技	Information Manufacturing	信息制造	Data Delivery	数据交付

Network Infrastructure	网络基础设施	Information Protection	信息保护	Data Application Services	数据应用服务
Network Information System	网络信息系统	Information Inclusion	信息纳入	Data Hosting	数据托管
Network Data Center	网络数据中心	Information Identification	信息识别	Data Consumption	数据消费
Network Facility	网络设施	Information Monitoring	信息监测	Data Property Rights	数据产权
Network Construction	网络建设	Information Recording	信息记录	Data Usage Rights	数据使用权
Network Storage	网络存储	Network Optimization	网络优化	Information Platform	信息平台
Network Governance	网络治理	Network Functions	网络功能	Information Sharing	信息共享
Network Access	网络接入	Network Applications	网络应用	Information Services	信息服务
Network Data	网络数据	Network Expansion	网络扩容	Information Consumption	信息消费
Network Technology	网络技术	Network Layout	网络布局	Information Resource Sharing	信息资源共享
Network Information Technology	网络信息技术	Network Operations	网络运营	Information Interconnection	信息互联互通
Network Information	网络信息	Network Running	网络运行	Information Service System	信息服务体系
Digital Network	网络数字	Data Engineering Technology	数据工程技术	Information Exchange	信息交换
Network Services	网络业务	Data Enterprises	数据企业	Information Service Level	信息服务水平
Network Provision	网络提供	Network Service Provider	网络服务提供者	Information Service Platform	信息服务平台

Network Services	网络服务	Cyber Sovereignty	网络空间主权	Online Trading	网络交易
Network Interconnection	网络互通	Data Security	数据安全	Online Sales	网络销售
Network Convergence	网络融合	Data Information Security	数据信息安全	Network Data Security	网络数据安全
Network Interoperability	网络互联互通	Data Evaluation	数据评价	Network Information Security	网络信息安全
Data Vendors	数据服务商	Data Service Providers	数据服务提供商	Network Security	网络安全性 / 网络安全
Data Interoperability	数据互联互通	Data Disclosure	数据公开	Network Risks	网络风险
Data Exchange	数据交易所	Data Supervision	数据监管	Data Supply	数据提供
Data Market	数据市场				

Table A6. Examples of GenAI-Related Dialogues on the IIPs

This table presents representative dialogues between listed firms and investors on China's Interactive Investor Platforms (IIPs) concerning generative AI adoption. The examples include one case where firms disclosed connections to foreign GenAI models (ChatGPT), one case involving domestic GenAI models (DeepSeek), and one case where firms explicitly stated non-adoption ("Not Deployed"). The dialogues illustrate the heterogeneity in firms' responses to investor inquiries about GenAI integration. In each example, the investor's adoption-focused question and the firm's reply are highlighted in bold for clarity.

Stock ID	Stock Name	Question Date	Reply Date	Investor Question	Firm Response
002045	Guoguang Electric	2023/5/29	2023/6/4	It is said that the company, together with international enterprises, has developed the world's first AI smart speaker based on ChatGPT, directly integrating ChatGPT. What is the progress of this project? When will it be released?	**The company ... has already released a video showcasing the ChatGPT-enabled smart speaker developed, designed, and produced by the company.** …
002010	Transfar Logistics	2025/2/11	2025/2/13	… Has it integrated DeepSeek?	... **Regarding the DeepSeek model, the company has recently completed localized deployment**, currently applied in campus customer service and dedicated line intelligent quotation scenarios…
000066	China Great Wall	2023/3/28	2023/3/31	Does the company have any plans in the field of OpenAI's ChatGPT? Has ChatGPT been integrated?	... **The company has not yet integrated ChatGPT…**