# Open-source Generative AI and Firm Value: Evidence from the Release of DeepSeek*

**Yue ZHAO**
Singapore Management University
yuezhao.2021@phdacc.smu.edu.sg

July 2025

**Abstract**

In this paper, I exploit the release of DeepSeek in January 2025 as an exogenous shock to the availability of open-source Generative AI (GenAI) models and investigate its value implications for corporate GenAI adopters. I document that U.S. firms with high pre-event GenAI exposure earn an average cumulative abnormal return of 1.2% over the event window, relative to low-exposure firms. The effect is stronger for firms with tighter financial constraints and for firms with proprietary information concerns, consistent with the benefits of open-source GenAI in reducing adoption costs and mitigating privacy concerns. In the post-event period, high-exposure firms are more likely to articulate adoption plans and benefits of open-source models during conference calls, and to embed DeepSeek in their algorithms shared on GitHub. High-exposure firms also experience an upward revision of analyst forecasts and a more positive media tone. In contrast, GenAI providers and their hardware suppliers experience negative abnormal returns. Lastly, the baseline analysis for Chinese firms yields a larger market impact of the DeepSeek's release. Overall, the findings indicate that open-source GenAI can further unlock the valuation potential of adopting GenAI.

**Keywords**: Generative AI, DeepSeek, Open-source, Proprietary Information, Firm Value

**JEL classifications**: G14, M41, O33

*Preliminary, Please Do Not Circulate Without Permission*

* I wish to thank my dissertation committee chair, Qiang Cheng, and dissertation committee members, Pengkai Lin, and Holly Yang for their invaluable guidance. I wish to thank the workshop participants at Singapore Management University for their helpful comments. I also acknowledge the support from the SMU Presidential Doctoral Fellowship. Please contact Yue Zhao at yuezhao.2021@phdacc.smu.edu.sg for comments.

"*As high-performing, low-cost, and openly available models proliferate, AI's accessibility and impact are set to expand even further*".

—AI Index Report 2025 by Stanford University[1]

## 1. Introduction

Generative artificial intelligence (GenAI) is expected to deliver significant benefits to business practices (Eisfeldt, Schubert, Taska, and Zhang 2024; Acemoglu 2025). Despite the widespread enthusiasm about GenAI, many firms struggle to capture the value of GenAI adoption and thus delay the adoption of or large-scale deployment of GenAI models due to various operational and strategic challenges (Boston Consulting Group [BCG] 2024; McKinsey & Company 2025). In particular, managers frequently cite high implementation costs and concerns about exposing proprietary information when using third-party *closed-source* GenAI tools such as ChatGPT (Bousquette 2025a; Cisco 2024). In contrast, *open-source* GenAI models can be downloaded and used freely, and deployed locally without sharing proprietary information with external GenAI providers. In this paper, I utilize the release of DeepSeek in January 2025 as an exogenous shock to the availability of open-source GenAI models, and examine whether firms with high pre-event GenAI exposure—the extent to which a firm's economic activities can be automated or augmented by GenAI—experience an increase in firm value after the DeepSeek rollout.[2,3]

DeepSeek represents a major advance among open-source GenAI models, with comparable performance to the leading OpenAI's model. Investigating the impact of the DeepSeek rollout can shed light on the ongoing debate about the extent to which GenAI can be used to generate benefits

---

[1] See https://hai.stanford.edu/ai-index/2025-ai-index-report.

[2] DeepSeek models are available under the permissive MIT license, a typical open-source license that grants users the right to use, study, modify, and distribute both the software and its source code freely. However, because the training data remain proprietary, technically speaking, the models do not satisfy the AI-community' stricter definition of "fully open-source," which requires both code and data to be freely available.

[3] I use DeepSeek, DeepSeek models, and DeepSeek-R1 model interchangeably.

for corporation (Acemoglu 2025). The open-source nature of DeepSeek differentiates it from the prevailing closed-source GenAI by attenuating two frictions that have constrained the value-creation potential of GenAI, i.e., implementation costs and privacy concerns. First, using closed-source GenAI tools have been expensive (Krause 2025a; Bousquette 2025a). Firm users often face steep pay-per-use fees or subscription costs for GenAI services. Fine-tuning a GenAI model can further require multi-million-dollar initial investments and ongoing expenses per user.[4] Some managers have postponed or cancelled at least one GenAI project because of widespread cost concerns.[5] Second, another hurdle is privacy concerns about using third-party GenAI models (Blankespoor, deHaan, and Li 2024). Integrating such GenAI models necessarily requires firms to send requests, which may involve sensitive information (e.g., product details, R&D materials, customer data), to external hosts, raising concerns that such information could be made public and accessed by competitors (Cisco 2024; Jia, Li, Xu, and Zhang 2025). Many firms have restricted employees' use of ChatGPT for work, citing fears of data leakage and loss of trade secrets.[6]

These two frictions associated with using closed-source GenAI models, such as ChatGPT, substantially reduce the incentives of firms with high GenAI exposure to adopt such models more broadly and intensively. Prior studies document the impact of ChatGPT adoption on various outcomes (e.g., Eisfeldt et al. 2024; Bertomeu, Lin, Liu, and Ni 2025a; Bradshaw, Ma, Yost, and Zou 2025; Brynjolfsson, Li and Raymond, 2025; Cheng, Lin, and Zhao 2025; Ecker, Li, Li, and Wu 2025), indicating the value-enhancing role of OpenAI's closed-source GenAI models. The findings of this line of research and the two concerns discussed above compel one to wonder whether corporate adopters will realize more benefits from adopting GenAI tools if OpenAI

[4] See https://www.techrepublic.com/article/30-generative-ai-projects-fail-amid-high-costs-and-risks/.
[5] See https://www.ibm.com/thought-leadership/institute-business-value/report/ceo-generative-ai/ceo-ai-cost-of-compute.
[6] See https://www.carusolawoffice.com/post/privacy-concerns-lead-major-companies-to-restrict-employee-use-of-chatgpt-for-work.

becomes truly "open" and releases an open-source GenAI model with comparable capabilities to ChatGPT. A lack of such open-source GenAI models remains a main challenge in investigating this question.[7]

I overcome this challenge by exploiting the release of DeepSeek's open-source GenAI models on January 20, 2025, a disruptive moment when an open-source GenAI model shocked the world (Banerji, Fitch, Langley 2025). DeepSeek's R1 model performs on par with OpenAI's leading model at a fraction of training costs (Gibney 2025). More importantly, unlike OpenAI, DeepSeek released the "weights" behind its model for the public to freely download, use, and customize. The open-source feature of DeepSeek significantly mitigates the two frictions discussed above, thereby delivering benefits to firms with high exposure to GenAI. First, firms can download DeepSeek models from GitHub or Hugging Face for free and locally utilize the models without paying recurring fees for API requests. Firms can also use DeepSeek's API services through cloud platforms like Azure, paying around 1/30 of the running costs of OpenAI's comparable model (o1) (Gibney 2025). Such cost reductions make it more feasible for high-exposure firms to experiment with and scale up GenAI solutions within their budget constraints. Second, because DeepSeek enables local or private-cloud deployment, firms no longer need to send requests that may include any sensitive data to third-party GenAI providers (Lin 2025). This local deployment and control can alleviate the proprietary concerns for high-exposure firms.

Collectively, with the DeepSeek's release mitigating both concerns, high-exposure firms are expected to adopt GenAI more broadly and intensively, further unlocking GenAI's productivity gains and value-creation potential. As such, relative to low-exposure firms, high-exposure firms

---

[7] Although Meta's Llama and Google's Gemma models are open to anyone to view, they are also not seen as being truly open source because the way users apply the models is restricted by licenses (See https://www.weforum.org/stories/2025/02/open-source-ai-innovation-deepseek/). Their performance is not comparable to OpenAI's closed-source models. Compared with earlier open-source GenAI releases, DeepSeek has triggered broader public discussions and a stronger market reaction, thereby providing a setting with greater test power.

will experience an increase in firm value after the DeepSeek's release. However, adopting open-source GenAI is not frictionless. It imposes switching costs on high-exposure firms that already adopted closed-source GenAI, and shifts hosting, maintenance, and cybersecurity expenses to the adopters. DeepSeek may deliver weaker domain accuracy because its training data may have limited coverage of U.S. contexts. It can also raise geopolitical concerns. Therefore, whether open-source GenAI can create more value for high-exposure firms remains an open question.

To investigate the research question, I conduct an event study that exploits the release of DeepSeek on January 20, 2025. The event window [0,5] begins on the release date and ends on January 28, when the market attention to DeepSeek reached the peak.[8,9] I focus on U.S. firms because the U.S. is currently the global leader in both developing cutting-edge GenAI models and deploying them in the enterprise, serving as a prime lens on GenAI's development. Since this paper focuses on the value implication of firms' adoption of GenAI, rather than the production of GenAI, I exclude GenAI providers and their hardware suppliers from the main sample.[10] Considering that the open-source GenAI models became universally accessible to all firms upon release, I exploit the years of 2023 and 2024, i.e., the period between ChatGPT's 2022 launch and DeepSeek's 2025 release, to construct a text-based measure of firms' exposure to GenAI based on the transcripts of earnings conference calls. I count the frequency of the GenAI-related keywords in each quarterly earnings conference call as the quarterly GenAI exposure measure. I then construct a firm-level GenAI exposure measure by taking average of quarterly measures across the two years for each firm. I convert the continuous measure into an indicator variable based on the sample median (i.e.,

[8] Please note that the actual day 0 is January 21, 2025 (Tuesday) since January 20 (Monday) is not a trading day.
[9] Figure 1 indicates that the market attention on "DeepSeek" (as proxied by Google Search Index in the U.S. region) jumps on the release date, climbs steadily, and peaks on January 28, 2025. DeepSeek overtook ChatGPT to become the most downloaded free app on the Apple App Store in the US on January 26 (Sunday) and set off a broad sell-off that knocked the S&P 500 down 1.5% and the Nasdaq 100 nearly 3% on January 27, 2025 (Reinicke 2025). The results are robust to using different event windows, as discussed in Section 4.2 Sensitivity Tests.
[10] Section 5.4 examines the value implications of the DeekSeek's release on GenAI providers and their hardware suppliers.

0) as the primary measure in the empirical analyses.[11] Because the GenAI exposure measure is constructed from pre-event data, it is unlikely to be influenced by the DeepSeek's release.

Using a sample of U.S. publicly traded firms, I examine cumulative abnormal returns based on the Fama-French three-factor model during the event window. The baseline results show that relative to low-exposure firms, high-exposure firms experience an increase in market value by around 1.2% following the DeepSeek release. A back-of-the-envelope estimate represents an increase of $200.4 billion market value of the 733 sample firms with high GenAI exposure, corroborating the view that the value of GenAI will "end up migrating up to essentially the application layer" (Rosenbush 2025).[12] Because general AI usage also create value as shown in existing studies (e.g., Chen and Srinivasan 2024; Babina, Fedyk, He, and Hodson 2024; Tucker, Wang, and Zhao 2025), I control for non-GenAI AI exposure throughout the analyses. Corroborating the main inference, firms with high AI exposure but low GenAI exposure do not experience an increase in firm value, consistent with the notion that it is specifically GenAI-driven capabilities that generate incremental firm value upon the release of DeepSeek.

I conduct several sensitivity tests to ensure the robustness of the main results. First, I find that high-exposure firms continue to experience a positive market reaction in the following two weeks (one month) from day 6 to day 15 (25), mitigating the concern on any return reversal driven by overreaction. Second, I use buy-and-hold abnormal return or CAR [0,5] based on the market model as the outcome variable and obtain the same inferences. Third, the main results are robust to both an extended event window and a three-day window ending on January 28, when the market attention to DeepSeek peaked. Fourth, I replace the indicator variable of GenAI exposure with a

[11] In sensitivity tests, I redefine the GenAI exposure measure in alternative ways (e.g., using the presentation session of conference calls, using SEC 10-K filings, using a continuous measure or a quartile measure), and obtain the same inferences.
[12] The economic magnitude is also comparable to Nvidia's record one-day loss of around $600 billion on January 20, 2025.

quartile rank variable or a continuous variable in logarithm and the results still hold. Fifth, I use the presentation session of earnings conference calls and SEC 10-K filings to re-construct two alternative measures of GenAI exposure, and obtain the same inference. Lastly, I conduct two placebo tests using tariffs-related keywords to construct a pseudo-GenAI exposure measure and using a pre-event period [-6,-1] as a pseudo-event window, and I do not find significant results.[13]

To further support the inferences and shed light on the benefits of open-source GenAI in reducing adoption costs and mitigating privacy concerns, I conduct two sets of cross-sectional analyses. First, financially constrained firms are more likely to have cost concerns about implementing closed-source GenAI models at high prices (e.g., Lerner and Schankerman 2013; Bertomeu, Lin, Liu, and Ni 2025b). I find that the main results are more pronounced for firms subject to tighter financial constraints, consistent with the cost-reduction benefit of open-source GenAI. Second, firms in competitive industries, firms with high R&D intensity, and firms providing business or personal services likely worry about sensitive data being inadvertently accessed by rivals and other stakeholders through the third-party GenAI platforms. I find that the main results are more pronounced for firms facing such proprietary information risks, consistent with the privacy-mitigation benefit of open-source GenAI.

While the return tests above provide early evidence on investors' perceived value of open-source GenAI, they cannot rule out the possibility that the market reaction is due to attention-driven GenAI hype. To examine whether investor perception is backed by firms' concrete actions and information intermediaries' responses, I conduct several follow-up analyses to provide additional insights. First, to validate that the documented value effect is at least partially attributed

[13] I use tariffs-related keywords because there were many discussions on tariffs since January 2025. Since the inauguration of Donald Trump as the 47th president of the United States also took place on January 20, 2025, I also use President-election-related keywords and the placebo results remain insignificant (untabulated).

to the benefits from adopting open-source GenAI, I use ChatGPT API to analyze the conference call transcripts in the post-DeepSeek period. I find that high-exposure firms are more likely to articulate their actual or planned adoption of DeepSeek and other open-source models developed by followers and to discuss the benefits of adopting such open-source GenAI models. Moreover, through manual searches of corporate GitHub accounts, I find that high-exposure firms are more likely to embed DeepSeek models into the algorithms they publish on GitHub.

Second, to corroborate the main findings, I explore changes in expectations of two information intermediaries in the capital market—financial analysts and business press. I find that after the DeepSeek release, high-exposure firms experience an upward revision of analyst forecast as well as a more positive tone in media reporting, relative to low-exposure firms.

Third, the valuation gains documented for GenAI adopters in the baseline analyses naturally raise a question that whether GenAI model providers suffer. DeepSeek may erode the economic rents of GenAI providers and their upstream hardware suppliers.[14] I document a significantly negative CAR [0,5] of -4.6% for these supply-side firms. Based on this sample, I find that firms that discussed GenAI in the pre-event period experience a more negative market reaction.

Lastly, I focus on a sample of Chinses listed firms to explore whether the effect is generalizable to the Chinese setting, given that DeepSeek is developed by a Chinese firm and ChatGPT has not been accessible to Chinese users. I find even stronger results for Chinese firms than for the U.S. sample. I also conduct a field experiment on the two major Chinese investor interactive platforms (IIPs) by soliciting firms' responses to a question regarding the deployment status and the intended business use of DeepSeek. The responses from 925 firms show that 23.2% (31.6%) of them have already deployed (plan to deploy) DeepSeek, and internal-management applications are mentioned

[14] A known example is that on January 27, 2025, Nvidia experienced a decline in stock price by 17% and lost approximately $600 billion in market capitalization (Reinicke 2025).

slightly more often than product- or service-oriented uses.

This paper contributes to existing literature in several ways. First, this paper contributes to the growing literature on GenAI adoption by shifting the focus from closed-source GenAI to open-source alternatives that further unleash GenAI's value-creation potential for corporate adopters. Prior studies exploit shocks related to ChatGPT (e.g., its release, outages, and bans) and document its favorable impact on various outcomes, including firm value (Eisfeldt et al. 2024; Bertomeu et al. 2025b), financial reporting (Blankespoor et al. 2025), customer-support services (Brynjolfssonet al. 2025), financial analysts' research (Bertomeu et al. 2025a; Bradshaw et al. 2025), investor information processing and trading activities (Cheng, Lin, and Zhao 2025; Ecker et al. 2025). However, the high implementation costs and privacy concerns of closed-source GenAI reduce the net benefits of GenAI adoption. This paper shows that open-source GenAI can ease both frictions and unlock valuation gains for GenAI adopters. This paper also responds to Eloundou et al.'s (2024) call for systematic tracking GenAI adoption and its real-world impacts.

Second, this paper contributes to the literature on the value of open-source software.[15] Existing studies mainly focus on firms that publicly disclose the source code of the software (Lerner and Tirole 2002; Lerner and Tirole 2005), showing that open-sourcing can create private value (Yang 2023; Emery, Lim, and Ye 2024; Coleman, Fronk, and Valentine 2025), improve innovation, productivity, and financial performance (Stam 2009; Nagle 2018), and facilitate external financing (Conti, Peukert, and Roche 2025) for the releasing firms. Nascent studies focusing on the demand-side adopters find that adopting open-source software can improve firm productivity (Nagle 2019) and help with investors' information processing (Cheng, Crowley, Lin, and Zhao 2025). This paper

[15] From an accounting perspective, releasing open-source software can be also regarded as voluntary disclosure of non-financial information (Coleman, Fronk, and Valentine 2025). Accordingly, the analysis of open-source GenAI also enriches the broader literature on innovation disclosure (e.g., Merkley 2014; Kim and Valentine 2021; Dyer et al. 2024).

complements the demand-side strand of literature by identifying a novel, GenAI-specific benefit of open-source software—local deployment that mitigates privacy concerns—which is largely absent from earlier open-source settings (Greenstein and Nagle 2014; Nagle 2019).

Third, the paper complements the broad literature on technology and AI adoption by firms. Recent studies examine the impact of AI adoption (e.g., Babina et al. 2024; Chen and Srinivasan 2024; Lem 2024; Awyong et al. 2024; Bai and Jin 2024; Chen et al. 2025; Tucker et al. 2025). GenAI is different from conventional AI because of its versatility, adaptability, and accessibility (Cheng, Lin, and Zhao 2025), and thereby has potential to create incremental value for firms. This paper contributes to the literature by demonstrating that, after controlling AI exposure, firms with high GenAI exposure still experience an increase in firm value after the release of DeepSeek.

## 2. Institutional Background and Hypothesis Development

### *2.1 Institutional Background*

#### 2.1.1 GenAI Adoption Before the Release of DeepSeek

Prior to the DeepSeek's release in 2025, U.S. firms that experimented and adopted GenAI models typically deployed them using closed-source GenAI models accessed directly through model vendors' APIs or through major cloud computing platforms like Microsoft Azure.[16] This approach provides immediate scalability and convenience but raises significant concerns around the recurring costs of using API services and the proprietary information risk that comes with sending sensitive data to third-party closed-source GenAI vendors.

First, implementation costs of GenAI have been significant (Krause 2025a; Bousquette 2025a),

[16] It is estimated that closed-source models accounted for 80% to 90% of the market in 2023, with the majority of share going to OpenAI (See https://a16z.com/generative-ai-enterprise-2024/). By the end of 2024, OpenAI was still the leading GenAI provider with the largest market share at around 75% (See https://firstpagesage.com/reports/top-generative-ai-chatbots/). As of April 2024, over 65% of the Fortune 500 use Azure OpenAI Service, and tens of thousands of other firms across industries and around the world are innovating with Azure AI, which provides cutting-edge APIs and models (See https://azure.microsoft.com/en-us/blog/microsoft-is-a-leader-in-the-2024-gartner-magic-quadrant-for-cloud-ai-developer-services/).

leaving GenAI adopters questioning the benefits of many GenAI pilots (Bousquette 2025b). Besides using GenAI models directly through API services, fine-tuning the model with firms' own data and using the fine-tuned model further incur a significant amount of recurring fees.[17] In an industry poll, nearly 75% of firms reported their cloud bills became "unmanageable" after adding GenAI workloads.[18] Such cost pressure explains why many firms postponed large-scale GenAI adoption although they believe in the long-run productivity upside.[19]

Second, proprietary information concerns have also slowed firms' adoption of GenAI at scale. Blankespoor et al. (2024) list privacy concerns about using closed-source GenAI among the risks and costs that may rode the net benefits of GenAI adoption. Deploying closed-source GenAI necessarily requires firms to send requests to external hosts. The user requests may involve product details, R&D materials, trade secrets, and customer information, raising concerns that such information could be shared with the public including competitors (Cisco 2024; Jia et al. 2025). For example, Samsung found that its employees uploaded sensitive internal source code to ChatGPT and thus it banned the use of GenAI on firm-owned devices and warned employees using GenAI on personal devices not to submit any proprietary information (Gurman 2023). Many other firms also have decided to restrict employee usage of third-party GenAI services.[20] The discussions above highlight that the expected proprietary costs can dominate the GenAI adoption calculus, such that even firms with high GenAI exposure may slow the pace of adoption in exchange for tighter control over proprietary information.

### 2.1.2 DeepSeek Rollout in January 2025

[17] Gartner estimates that customizing a top-tier language model could cost a company $5–20 million upfront plus $8k–$21k per user annually (See https://www.techrepublic.com/article/30-generative-ai-projects-fail-amid-high-costs-and-risks/)
[18] See https://www.supplychaindive.com/news/generative-ai-drives-unmanageable-cloud-cost-finops/729246/.
[19] See https://www.ibm.com/thought-leadership/institute-business-value/report/ceo-generative-ai/ceo-ai-cost-of-compute.
[20] See https://www.carusolawoffice.com/post/privacy-concerns-lead-major-companies-to-restrict-employee-use-of-chatgpt-for-work. Moreover, reliability of GenAI responses may present additional hurdles (Blankespoor et al. 2025; Acemoglu 2025), prompting enterprises to maintain human oversight and cautious scaling strategies to ensure responsible GenAI use.

OpenAI has established itself as the GenAI industry leader since the debut of ChatGPT in November 2022. DeepSeek, a Chinese firm, dramatically shifted the GenAI landscape with the release of its DeepSeek-R1 model in January 2025. DeepSeek-R1was developed on less advanced chips at a fraction of training costs, approximately $6 million in comparison to around $78 million for OpenAI's GPT-4 model. In contrast, DeepSeek-R1 has comparable performance to the leading closed-source GenAI models like OpenAI's o1 (Gibney 2025). The release of DeepSeek challenges a prevailing assumption that only big tech firms with massive computing power and advanced chips can lead in the era of (Gen)AI, sparking a broad stock selloff on January 27, 2025.

DeepSeek has made its GenAI models publicly available under an open-source license, allowing all firms to download the model weights freely for local deployment. A series of DeepSeek models, including DeepSeek-R1 (617B), DeepSeek-R1-Zero (617B), and six distilled models (1.5B, 7B, 8B, 14B, 32B, 70B) based on Qwen and Llama provide adopters with a menu of model sizes and capabilities, enabling them to select the configuration that best fits their computational budgets and task requirements.[21] In addition, the DeepSeek models can be hosted on-premises or in private-cloud environments.[22] All prompts and outputs therefore remain behind the firm's firewall, mitigating the privacy concerns associated with closed-source GenAI models.

The initial hardware investment to deploy DeepSeek locally is also not costly. The innovative Mixture-of-Experts (MoE) architecture used by DeepSeek selectively activates only the relevant sub-models needed for each query and thus significantly enhances computational efficiency. Rather than executing all 671 billion parameters of DeepSeek-R1 on each inference pass, the model activates only a subset of approximately 37 billion parameters in real time according to the

---

[21] DeepSeek is compatible with standard PyTorch and popular inference stacks such as vLLM, thus firms can incorporate the model into existing ML-Ops pipelines, mitigating the interoperability concern associated with open-source software (Lerner and Schankerman 2013).

[22] See https://github.com/deepseek-ai/DeepSeek-R1/blob/main/DeepSeek_R1.pdf for technical details of DeepSeek-R1 model.

semantic content of the input. This selective routing not only slashes upfront capital expenditure on GPUs but also reduces recurring inference costs for GenAI adopters. According to estimates, running the quantized 671B DeepSeek-R1 model requires roughly $100,000 to $120,000 in GPUs.[23] Firms can leverage existing server setups or moderate hardware investments to deploy the model, rather than spending millions on new supercomputing clusters. Moreover, distilled variants, such as DeepSeek-R1-Distill-Qwen-7B, require even less hardware.

Besides local deployment, firms can also use DeepSeek models directly through its API services. In addition, major American cloud service providers such as Amazon Web Services (AWS), Microsoft, and Google also deploy DeepSeek's open-source model into their cloud computing platforms and offer DeepSeek models to broad users (Lin, Chin, and Huang 2025).

### *2.2 Hypothesis Development*

Firm with high GenAI exposure—the extent to which a firm's economic activities can be automated or augmented by GenAI—are more likely to deploy GenAI for automation and task complementarity purposes (Acemoglu and Restrepo 2018; Acemoglu 2025). Prior studies suggest that OpenAI's closed-source GenAI models are expected to deliver considerable benefits for high-exposure firms (e.g., Eisfeldt et al. 2024; Bertomeu et al. 2025b). Nevertheless, high implementation costs and proprietary information concerns inherent in closed-source GenAI substantially erode these net benefits, prompting managers to postpone large-scale deployment of GenAI models. Consequently, the modest macroeconomic effects of GenAI projected by Acemoglu (2025) will likely become even smaller, underscoring the need to explore alternatives that can unlock GenAI's potential by mitigating the dual frictions.

---

[23] See https://apxml.com/posts/gpu-requirements-deepseek-r1, https://www.reddit.com/r/selfhosted/comments/1ibl5wr/comment/m9j6m1e/, and https://medium.com/data-science-in-your-pocket/how-much-it-costs-to-run-deepseek-r1-locally-651f10006b95 for the estimates.

The release of DeepSeek's open-source GenAI models on January 20, 2025, can lead to a sharp reduction in both barriers. First, open-source GenAI models significantly cut the implementation costs for high-exposure firms who are more likely to benefit from adopting GenAI. One salient benefit of using nonpecuniary open-source software is that it is free (Lerner and Schankerman 2013; Nagle 2019). DeepSeek provides its models openly under permissive licenses, charging no licensing fees or revenue share for downloading and using the model. This contrasts with closed-source GenAI models that charge recurring fees of API requests and subscription fees. Firms can download the models for free and deploy the models locally or in their private cloud environments. Besides, the MoE architecture and other optimizations of DeepSeek drastically improve computational efficiency, thereby requiring fewer GPUs for local deployment and reducing the inference cost of each request from users.[24] As such, DeepSeek significantly cut the deployment costs without sacrificing inference performance, therefore raises the net present value of adopting GenAI.

The impact of DeepSeek's release is not necessarily directly from firms' adopting DeepSeek models. The availability of such open-source models also represents a technology shock to the entire GenAI landscape. Because the models are fully open-source, other GenAI providers can incorporate DeepSeek's training recipes for their own training and customizing. Several GenAI providers and cloud platforms, such as Microsoft's Azure and Amazon's AWS, have downloaded DeepSeek and added its models to their GenAI platforms. In addition, DeepSeek places intense downward pressure on industry pricing because it delivers performance on par with OpenAI's flagship o1 model yet offers API access at roughly 1/30 of o1's price. In the weeks after the release

[24] To illustrate its lower inference cost over closed-source GenAI, DeepSeek-R1's API is priced at only $2.19 ($0.55) per million output (input) tokens, in comparison to $60 ($15) per million output (input) tokens for a comparable OpenAI model (See https://www.techtarget.com/whatis/feature/DeepSeek-explained-Everything-you-need-to-know). Firms can also use DeepSeek's API provided by DeepSeek or popular cloud platforms like Azure, paying around 1/30 of the running costs of OpenAI's comparable model (o1) (Gibney 2025).

of DeepSeek, other major tech firms (e.g., OpenAI, Alibaba, Baidu, Tencent) cut GenAI service prices to stay competitive.[25]

Second, open-source GenAI can mitigate privacy concerns by allowing firms to use the model in their own secure environment. Firms can deploy DeepSeek models on-premises or in a private cloud, behind their firewalls as well as under their governance. Firms do not have to send sensitive data to outside GenAI vendors, and all data processing stays within the firm's infrastructure. For example, a bank (or a healthcare firm) could run DeepSeek on in-house servers so that customer (or patient) information will never leave their controlled environment. Additionally, firms can customize the DeepSeek models locally based on their internal data, without relying on fine-tuning services provided by external GenAI vendors. As such, the ability to run and fine-tune the model in house mitigates potential losses from data leakage and thus improves GenAI' value gains for high-exposure firms.

With open-source GenAI models reducing both the deployment costs and privacy concerns, high-exposure firms are expected to scale GenAI adoption for automation and task complementarity (Acemoglu 2025), and hence capture a larger incremental value unlocked by the release of DeepSeek. I therefore predict that the stock market reaction to the release of DeepSeek's open-source GenAI models will increase with a firm's pre-event GenAI exposure. I state the main hypothesis as follows:

***H1****: Firms with high GenAI exposure experience more positive cumulative abnormal returns after the release of DeepSeek models, relative to low-exposure firms.*

There are also reasons that I may not find results consistent with H1. First, the costs of

[25] For example, OpenAI released an affordable and cost-efficient model, "o3-mini" on January 31, 2025 (See https://openai.com/index/openai-o3-mini/). Chinese tech leaders also "flooded the market with a rapid succession of low-cost AI services, , undercutting premium offerings from the likes of OpenAI and Alphabet Inc.'s Google" (Rai and Sun 2025).

switching from closed-source to open-source models can be non-trivial for high-exposure firms, especially for those who have already embraced the closed-source software (Lerner and Schankerman 2013). Migrating fine-tuned workflows and rewriting application interfaces can be costly and time-consuming, such that the price discount of DeepSeek may not translate into net savings. Second, although open-source models present high-exposure firms with a considerably lower upfront cost, operating an open-source model shifts hosting, maintenance, and cybersecurity costs from the third-party vendor to the corporate adopter (Lerner and Schankerman 2013; Nagle 2019). This could offset much of DeepSeek's price advantage. Third, DeepSeek' training data, which have not been made publicly available (Gibney 2025), may have limited coverage of U.S. contexts and English-language materials. Lastly, although DeepSeek can be deployed locally, U.S. firms may hesitate to use a Chinese GenAI model out of geopolitical or national-security concerns.

## 3. Empirical Design

### *3.1 Data and Sample*

The sample construction begins with 5,847 firm-level observations with financial data from Compustat. I remove 845 observations with missing cumulative abnormal returns data, followed by 1,211 observations without conference call transcripts during 2023 and 2024. I then exclude 98 observations with missing control variables. Finally, to focus on GenAI adopters rather than producers, I exclude 219 observations from firms that directly develop foundational GenAI models or firms in the "Computers" and "Electronic Chips" industries under the Fama–French 48 industry classification. These sequential steps yield a final sample of 3,474 firm-level observations. Table 1 summarizes the sample selection process.

I obtain conference call transcripts data from the Capital IQ database. I collect daily stock prices from the Center for Research in Securities Prices (CRSP), firms' accounting information from Compustat, SEC filing data from the SEC EDGAR system through SEC API (https://sec-api.io/), news coverage data from RavenPack, and analyst forecasts data from the Institutional Brokers' Estimate System (I/B/E/S). I collect stock price and accounting data of Chinese firms from China Stock Market & Accounting Research (CSMAR) database.

### *3.2 Measurement of GenAI Exposure*

Since DeepSeek's open-source GenAI models became freely available to all public firms immediately after its release, it is *ex ante* unclear which firm would derive more incremental benefits from GenAI adoption. I argue that a firm's *ex ante* readiness to capture value from DeepSeek hinges on its managerial attention and engagement with GenAI before the release of DeepSeek. I follow a growing literature that infers technological exposure from firms' disclosure that targets capital market participants (e.g., Chen and Srinivasan 2024; Tucker et al. 2025).

To operationalize GenAI exposure, I construct a text-based measure using the full set of quarterly earnings conference call transcripts from Q1 2023 through Q4 2024.[26] This period is after ChatGPT's November-2022 debut so that early (potential) adopters have assimilated baseline GenAI capabilities (realized the benefits of GenAI), yet precedes DeepSeek's January-2025 rollout, ensuring that observed discourse is not contaminated by anticipation effects specific to DeepSeek. For each transcript, I compute the raw frequency of GenAI-related expressions appearing in presentation and Q&A sessions. The keyword dictionary is adapted from Jia et al. (2025), who assemble a comprehensive list of 41 terms that capture GenAI vendor (e.g., "OpenAI," "Claude,"

[26] The text-based measure complements the predominantly labor-focused measure employed in prior work (e.g., Eloundou et al. 2024; Eisfeldt et al. 2024; Bertomeu et al. 2025b). The labor-focused measure gauges the technological feasibility of task substitution or complementarity and thus captures the single-dimensional labor exposure to GenAI, which may overstate exposure when adoption frictions are high. In contrast, the text-based measure could aggregate the myriad considerations managers weigh when evaluating GenAI adoption (Jia et al. 2025). By relying on what firms communicated to capital-market participants, the measure embeds managerial readiness and willingness to adopt GenAI, thereby aligning more closely with information sets that underpin investors' cash-flow expectations and disclosure-based valuation models. Please also see Eloundou et al. (2024) for detailed discussions on limitations of the labor-exposure measure.

"Anthropic"), GenAI product (e.g., "ChatGPT," "Copilot"), and GenAI technology (e.g., "large language model," "Generative AI").[27] Appendix B presents several examples of GenAI-related discussions in conference calls.

I then average the quarterly GenAI exposure measure across the eight quarters in 2023-2024 to obtain a firm-level measure (i.e., the raw GenAI exposure measure). Because the distribution of this continuous variable is highly right-skewed, and also for ease of interpretation, I convert this continuous measure based on the sample median (i.e., 0) into an indicator variable (*GenAI Exposure*), which equals one if a firm discussed GenAI-related topics at least once during the conference calls in 2023 and 2024, and zero otherwise.[28] Nonetheless, the baseline results are robust to alternative measures of *GenAI Exposure*, including a quartile rank variable as in Chen and Srinivasan (2024), and a continuous variable in logarithm.

To differentiate the effect of GenAI from firms' exposure to traditional AI and digital technologies, I also construct *AI Exposure* based on the non-GenAI AI keywords listed in Chen and Srinivasan (2024). *AI Exposure* is defined similarly as *GenAI Exposure*, except that it takes a value of one for firms only discussing non-GenAI AI technologies (mentioning AI keywords but not mentioning GenAI keywords), and zero otherwise. If the valuation response to the release of DeepSeek is specific to GenAI rather than to conventional AI in general, firms with non-zero *AI Exposure* (so automatically with zero *GenAI Exposure* by definition) should not experience an increase in firm value. I therefore control for *AI Exposure* throughout the analyses to ensure the documented effects are driven by GenAI exposure specifically, rather than the exposure to conventional AI and digital technologies.

---

[27] I drop several keywords (e.g., Bard, Hugging Face, Transformer) from the list since they appear as manager names or general AI terms rather than GenAI-specific terms in the transcripts, thus introducing measurement errors into the GenAI exposure measure. The results still hold if I use the keywords list in Jia et al. (2025) (untabulated).

[28] Because I control for industry fixed effects throughout the analyses, the construction of the indicator variable is not based on industry median.

In addition, there is a possibility that low GenAI exposure may capture late adopters of GenAI, and DeepSeek's open-source release could plausibly create greater upside for them by lowering implementation costs and mitigating privacy concerns. If so, firms with high *GenAI Exposure* could receive a muted or even negative reaction, while firms with high *AI Exposure* (but with zero *GenAI Exposure*) would experience positive abnormal returns because investors expect them to leverage their existing AI capabilities to incorporate DeepSeek. Including *AI Exposure* as a control thus could also rule out that *GenAI Exposure* simply captures early versus late adopters.

### *3.3 Empirical Design*

I use the following ordinary least squares (OLS) regression model to examine the value implication of the DeepSeek's release:

$$CAR[0,5]_i = \beta_0 + \beta_1 GenAI\ Exposure_i + Controls_i + \theta_{ind} + \varepsilon_i \quad (1)$$

where *CAR [0,5]* is the cumulative abnormal return (CAR) based on the Fama-French three-factor model during day 0 to day 5, with day 0 being January 21, 2025 (Deepseek's release date, January 20 was not a trading day), and with day 5 being January 28. Figure 1 indicates that the Google Search Index for "DeepSeek" in the U.S. region jumps on the release date, climbs steadily, and peaks on January 28, 2025.[29] *GenAI Exposure* is an indicator variable that equals one if firm *i* discussed GenAI-related topics at least once during the conference calls in 2023 and 2024, and zero otherwise. The coefficient on *GenAI Exposure*, $\beta_1$, captures the difference in CAR between firms with high GenAI exposure and those with low GenAI exposure. H1 implies a positive coefficient on *GenAI Exposure*.

I include a set of firm characteristics that could affect valuation dynamics: the logarithm of

[29] Although the Google Search Index for "DeepSeek" has declined since January 29, 2025, this does not necessarily indicate that DeepSeek's real-world impact is fading. Figure OA-1 shows that the number of DeepSeek downloads from Hugging Face platform continue to rise steadily, signaling sustained and growing adoption.

market capitalization (*Size*), the logarithm of firm age measured from the first Compustat appearance (*Age*), leverage ratio (*Lev*), return on assets (*ROA*), a loss indicator (*Loss*), book-to-market ratio (*BM*), asset tangibility (*Tangibility*), R&D intensity (*R&D*), and advertising intensity (*Advertising*). To mitigate the concern that the results are driven by exposure to broad AI technologies instead of exposure to GenAI specifically, I also control for *AI Exposure* throughout the analyses.[30,31] Appendix A provides detailed variable definitions. All continuous variables are winsorized at the 1st and 99th percentiles. I also include Fama–French 48 industry fixed effects to absorb time-invariant industry-specific influences. The t-statistics are based on robust standard errors as in Ishii and Xuan (2014).

### *3.4 Descriptive Statistics*

Panel A of Table 2 reports summary statistics for the variables employed in the baseline analyses. The average *CAR [0,5]* is around 0.2% with a standard deviation of 0.077. The mean of *GenAI Exposure* is 0.211, indicating that most 21.1% of sample firms made GenAI-related discussions at least once in 2023–2024. I use OpenAI's API to analyze these discussions and find that around 90% of high-exposure firms explicitly mention they have already adopted GenAI. Turning to firm fundamentals, the average *Size* is 14.028, equivalent to around $1.2 billion; the average *Age* is 2.879 (roughly 18 years), and leverage averages 62.2%; *ROA* is modestly negative on average, with mean *ROA* of –0.076 and 39.0% of observations reporting a loss. The mean of

---

[30] I conduct determinant analyses to verify that *GenAI Exposure* reflects systematic firm traits. Table OA-1 shows that larger firms, younger firms, and firms with higher profitability, more R&D and advertising outlays, and lower asset tangibility are systematically more inclined to discuss GenAI even after adding Fama-French 48 industry fixed effects. The results on *Size*, *Age*, *R&D*, *Advertising* are consistent with the findings in prior studies (e.g., Jia et al. 2025; Chen and Srinivasan 2024). The adjusted $R^2$ triples from 0.112 to 0.338 after adding controls and Fama–French 48 dummies, indicating that these characteristics explain a meaningful share of the cross-sectional variation in *GenAI Exposure*. Note that I control these determinants and industry fixed effects throughout the analyses.

[31] I don't have a prior on the coefficient on *AI Exposure*. The coefficient can be indistinguishable from zero since the DeepSeek shock is specific to GenAI and thus might not affect firms with only traditional AI adoption. The coefficient can also be negative if investors discount firms sticking using traditional AI in the coming era of GenAI. The coefficient can be positive as well if investors assume synergy between traditional AI and GenAI.

*BM* is 0.654, tangible assets represent 22.2% of total assets, and R&D and advertising intensities average 6.8% and 1.3% of operating expenses, respectively, with medians of zero for both. The descriptive statistics of the control variables are largely comparable to those reported in prior studies (e.g., Eisfeldt et al. 2024).

Panel B of Table 2 reports the correlations between the variables used in the main analyses. The correlation coefficient between *CAR [0,5]* and *GenAI Exposure* is 0.121 and statistically significant at 1% level, providing univariate evidence that firms with high GenAI exposure ahead of the DeepSeek's release experience larger valuation gains during the event window than other firms. Turning to potential multicollinearity, the correlation coefficient with the highest absolute value in the matrix is –0.638 (between *ROA* and *R&D*). Untabulated results from the variance inflation factor analysis indicate that multicollinearity is not a concern.

Panel C of Table 2 shows that *GenAI Exposure* is highly concentrated in information- and service-oriented sectors. Nearly four-fifths of Printing & Publishing firms (80.0%) and two-thirds of Business Services firms (67.3%) are exposed to GenAI, whereas traditional manufacturing and resource industries, e.g., Agriculture, Beer & Liquor, and Coal, have little exposure. Most other sectors fall between 10% and 30%, with financials (Banking 14.1%, Insurance 18.1%) and retail-facing industries (Retail 24.7%, Restaurants, Hotels, Motels 21.1%) falling near the upper end. These patterns indicate a significant variation of *GenAI Exposure* across industries.

## 4. Main Analyses

### *4.1 Baseline analyses*

In this section, I examine the stock market reaction to the release of DeepSeek during the [0,5] event window. Table 3 presents the results of baseline analyses. Column (1) controls for *AI Exposure* as a benchmark, Column (2) includes other firm-level controls, and Column (3) further introduces Fama–French 48 industry fixed effects. The coefficient on *GenAI Exposure* is positive and statistically significant at 1% level in all three columns ($t$ = 6.80, $t$ = 3.02, and $t$ = 2.63, respectively). The effect is economically meaningful; firms with high pre-event GenAI exposure earn around 1.2% higher abnormal returns during the event period compared with low-exposure firms, consistent with H1. Scaling this return by the combined market capitalization of the 733 high-exposure firms in the sample (≈ $16.7 trillion) indicates an aggregate shareholder-wealth creation on the order of $200.4 billion.

In contrast, the coefficient on *AI Exposure* is not significant, mitigating the concern that the documented valuation impact of DeepSeek is driven by firms' exposure to conventional AI. In addition, the positive coefficient on *GenAI Exposure*, together with insignificant coefficient on *AI Exposure*, are inconsistent with the alternative interpretation that the GenAI exposure measure simply captures early versus late adopter.

### *4.2 Sensitivity Tests*

I conduct a series of sensitivity tests to ensure the robustness of the baseline results. First, to mitigate the concern that the main results simply capture an overreaction to the DeepSeek's release in a short-term window and there might be a reversal in the long run, I examine whether high-exposure firms continue to experience a more positive market reaction in the following two weeks (one month) from day 6 to day 15 (25).[32] Panel A of Table 4 presents the results using *CAR [6,15]* and *CAR [6,25]* as the dependent variables. The coefficient on *GenAI Exposure* remains positive and statistically significant in both columns ($t$ = 1.96 and 2.76, respectively), indicating that the valuation gains for high-exposure firms persist for at least one month after the release of DeepSeek.

Second, to ensure that the baseline results are not sensitive to how abnormal return is

[32] While the longer event windows might be subject to contamination by other related events, it allows investors to take more time to digest and incorporate the value implication information into stock prices.

calculated, I use buy-and-hold abnormal returns over the [0,5] window (*BHAR [0,5]*) and CAR [0,5] based on the market model (*CAR [0,5]_Market_Adj*) as alternative outcome variables. Panel B of Table 4 presents the results. The coefficient on *GenAI Exposure* remains positive and statistically significant in both columns (t = 2.58 and 3.06, respectively).

Third, I use [-1,6] and [3,5] as two alternative windows to ensure that the main results are not sensitive to the extended event window or the window ending on January 28, when the market attention to DeepSeek peaked. Panel C of Table 4 presents the results using *CAR [-1,6]* and *CAR [3,5]* as dependent variables. The coefficient on *GenAI Exposure* remains positive and statistically significant in both columns (*t* = 2.69 and 2.22, respectively).[33]

Fourth, in the baseline analyses, I use an indicator variable (*GenAI Exposure*), because it is easy to interpret the coefficient and mitigates influence from "super-talkers." However, the conversion may discard meaningful variation in the intensity of GenAI exposure. To address this concern, I replace the indicator variable with a quartile rank variable as in Chen and Srinivasan (2024), and a continuous measure in logarithm. As shown in Panel D of Table 4, the coefficient is positive and statistically significant in both columns (t = 3.64 and 3.75, respectively).

Fifth, to ensure the baseline results are not sensitive to the source of GenAI discussions, I use the presentation session of earnings conference calls and SEC 10-K filings to re-construct two alternative measures of GenAI exposure.[34] As reported in Panel E of Table 4, the coefficient on *GenAI Exposure* remains positive and significant in both columns (t = 3.17 and 2.27, respectively).

Finally, I conduct two placebo tests using tariffs-related keywords to construct a pseudo-GenAI exposure measure and using a pre-event period [-6,-1] as a pseudo-event window. Panel F

[33] The results are also robust to several alternative windows including [-2,7], [0,6], [-1,5] (untabulated).
[34] The results are also robust to using the Q&A session of earning conference call transcripts to construct GenAI exposure (untabulated).

of Table 4 presents the results. In both placebo tests, the coefficient on *GenAI Exposure* is indistinguishable from zero (t = 0.14 and -1.36, respectively), indicating that the documented return premium is unique to GenAI exposure and to the actual DeepSeek rollout date.

Collectively, the sensitivity tests show that the baseline results are robust to alternative event windows, abnormal return metrics, and exposure definitions. The findings reinforce the conclusion that firms with high GenAI exposure experienced economically meaningful valuation gains after the rollout of DeepSeek.

### *4.3 Cross-sectional Analyses Based on Financial Constraints*

The baseline results suggest that once DeepSeek's open-source GenAI models become available, firms with high GenAI exposure earn larger abnormal returns. To further support the inferences and shed light on the benefits of open-source GenAI in reducing implementation costs and proprietary information concerns, I conduct two sets of cross-sectional analyses.

First, if open-source GenAI can increase firm value by reducing the implementation costs, I would expect that the value impact of DeepSeek documented in the baseline results is stronger when firms are more constrained in their ability to fund large-scale GenAI projects. I rely on three complementary proxies that the prior literature used to capture financial constraints. Smaller firms and younger firms are more likely to be subject to financial constraints (e.g., Li 2011; Blankespoor et al. 2024; Bertomeu et al. 2025b). I also utilize a measure developed by Linn and Weagley (2024), who extend the coverage of the text-based financial constraint measures in Hoberg and Maksimovic (2015) by training a random forest model on accounting variables only.[35] I conduct cross-sectional analyses on financial constraints using the following regression:

$$CAR[0,5]_i = \beta_0 + \beta_1 GenAI\ Exposure_i \times CSVar_i + \beta_2 GenAI\ Exposure_i \quad (2)$$

[35] Please find the data at https://www.danielweagley.com/data.

$$+ \beta_3 CSVar_i + Controls_i + \theta_{ind} + \varepsilon_i$$

where *CSVar* is an indicator variable based on each of the proxy for financial constraints. Specifically, *Size_Small* is an indicator variable that equals one if *Size* (i.e., log market capitalization) is below the sample median, and zero otherwise. *Age_Low* is an indicator variable that equals one if *Age* (i.e., years since first Compustat appearance) is below the sample median, and zero otherwise. *Fin_Constrain_High* is an indicator variable if the financial constraints measure developed by Linn and Weagley (2024) is above the sample median, and zero otherwise.

Panel A of Table 5 presents the results. The coefficient of the interaction term is positive and statistically significant in all columns (t = 1.76, 3.36, and 3.03, respectively). These results indicate a stronger effect of the DeepSeek's release on firms with tighter financial constraints, consistent with the notion that the open-source GenAI model can further unlock the valuation potential of GenAI adoption for high-exposure firms by slashing the deployment costs.

*4.4 Cross-sectional Analyses Based on Proprietary Concerns*

Second, if open-source GenAI can improve firm value by alleviating proprietary information concerns, I would expect the baseline effect to be stronger for firms that are more likely to have privacy concerns. I draw on three proxies commonly used to capture proprietary information concerns. Prior studies suggest that the expected cost of proprietary information leakage rises with more intensive competition (Verrecchia and Weber, 2006; Dedman and Lennox 2009), and when firms' competitive advantage rests on proprietary information such as R&D information and trade secrets (Glaeser 2018; Koh and Reeb 2015). Additionally, firms in business services and personal services have more customer data and thus face higher data leakage risks.

To test the privacy-mitigation benefits of open-source GenAI, I use the same regression model as Equation (2) except that *CSVar* is one of the three proprietary concern measures.

*Competition_High* is an indicator variable that equals one if the product similarity score developed by Hoberg and Phillips (2016) is above the sample median, and zero otherwise. *R&D_High* is an indicator variable that equals one if the R&D expenses scaled by operating expenses is above the sample median, and zero otherwise.[36] *Service_Ind* is an indicator variable that equals one if firm *i* is in the business or personal services industry as defined in Fama-French 48 industry classification, and zero otherwise.

Panel B of Table 5 reports the results from cross-sectional analyses on proprietary concerns. The coefficient of the interaction term remains positive and statistically significant in all columns (t = 3.38, 2.81, and 3.99, respectively). The findings indicate a larger effect of the DeepSeek's release for firms with more intensive product market competition, firms with higher R&D intensity, and firms in business or personal service industries, consistent with the notion that DeepSeek's open-source deployment mitigates privacy concerns and thus generate valuation gains for high-exposure firms.

## 5. Additional Analyses

Although the return tests above provide early evidence on investors' perceived value of open-source GenAI added to corporate adopter, I acknowledge that these results cannot rule out the possibility that the observed price reaction is simply attention-driven hype. To examine whether investor perception is backed by firms' concrete actions and information intermediaries' responses, I conduct several complementary analyses to provide additional support. I also examine whether GenAI providers and their hardware suppliers face adverse effects and investigate whether the baseline results can be generalized to Chinese firms.

[36] The results still hold for *Trade Secret*, which is an indicator variable that equals one if firm *i* mentioned "trade secret" or "trade secrecy" in Item 1 Business or in Item 1A Risk Factors in 10-K filings, and zero otherwise (Glaeser 2018).

### *5.1 Conference-Call Evidence on Adoption (Plan) and Anticipated Benefits*

First, to corroborate that the valuation effects documented above, I examine whether firms with high GenAI exposure become more vocal about adopting open-source models and about the benefits they expect to realize from GenAI adoption. I use GPT API to analyze the conference call transcripts after the end of the event window (January 28, 2025) till Aril 15, 2025, and identify whether managers of firm *i* mention that they (will) adopt DeepSeek or open-source models developed by followers (*DeepSeek Adoption*), and whether they discuss the benefit of adopting such models (*DeepSeek Benefit*).[37] I use the same regression as Equation (1) except that I replace the outcome variable with one of the two variables.

Table 6 reports the results. The coefficient on *GenAI Exposure* is positive and statistically significant in both columns (t = 8.16, and 7.41, respectively), indicating that firms with high GenAI exposure are more likely to mention that they (will) adopt GenAI and discuss the benefits of open-source GenAI after the release of DeepSeek. The evidence therefore corroborates the interpretation that the open-source GenAI at least spur real adoption intentions.

### *5.2 Use of DeepSeek Models in Firms' Algorithms Published on GitHub*

To test whether high-exposure firms actually use DeepSeek rather than merely mention it during the post-event conference calls, I manually search the sample firms' GitHub repositories for the keyword "DeepSeek". I construct an indicator variable, *DeepSeek GitHub*, that equals one if firm *i* has embedded DeepSeek models in their open-source algorithms shared publicly on GitHub, and zero otherwise. I use the same regression as Equation (1) except that I replace the outcome variable with *DeepSeek GitHub*.

[37] Since there were no DeepSeek-related discussions before the rollout of DeepSeek, the two outcome variables essentially capture the changes from pre- to post-DeepSeek periods, making Equation (3) a DiD analysis in spirit. The mean (standard deviation) of *DeepSeek Adoption* is 0.059 (0.344), and the mean (standard deviation) of *DeepSeek Benefit* is 0.051 (0.327).

Table 7 reports the results. The coefficient on *GenAI Exposure* is positive and statistically significant (t = 4.68), indicating that firms with high GenAI exposure are more likely to incorporate DeepSeek's open-source model into the code they host on GitHub after the release of DeepSeek. The evidence therefore at least helps alleviate the concern that high-exposure firms are simply doing GenAI washing by talking during the conference calls. Moreover, since GitHub activities can create private value for the releasing firms (Emery et al. 2024; Coleman et al. 2025), the findings suggest a channel through which DeepSeek may enhance the value of GenAI adopters.

### *5.3 Changes in Expectations of Information Intermediaries*

If investors rationally incorporate DeepSeek's implications into stock prices, the two main information intermediaries, i.e., financial analysts and business press, should also update their beliefs that shape market beliefs. Accordingly, I examine whether high-exposure firms experience more positive post-event revisions in analyst earnings forecasts and media reporting tone. For each firm, I measure the change in analyst forecasts (*Analyst_Forecast_Chg*) as the percentage change in analyst forecast consensus from pre-event to post-event periods. I similarly define the change in media sentiment (*Media_Sentiment_Chg*) using the news article sentiment provided by RavenPack database. I use the same regression as Equation (1) except that I replace the outcome variable with one of these change measures, respectively.

Table 8 reports the results. The coefficient on *GenAI Exposure* is positive and statistically significant in both columns (t = 1.78, and 1.78, respectively), suggesting that financial analysts and media broadly interpret DeepSeek as enhancing the value of GenAI adoption for high-exposure firms. The findings lend credibility to the argument that the market reaction to the DeepSeek rollout reflects informed reassessment of firm value rather than transient trading noise.

### *5.4 Evidence from GenAI Providers and Their Hardware Suppliers*

The valuation gains documented for adopters in the baseline analyses naturally raise the question of whether the existing GenAI providers and their upstream hardware suppliers suffer. DeepSeek may erode the economic rents of these firms whose revenues depend on large-scale GenAI training and development demand. There are two reasons why these firms might react negatively. First, DeepSeek disrupted a prevailing notion that "the bigger, the better". If DeepSeek's success suggests that cheaper training can achieve comparable performance, existing GenAI providers may lose the competitive edge (Krause 2025b). Second, because the MoE architecture can help deliver inference at a fraction of the hardware investments, suppliers of advanced chips could see a contraction in future orders. However, since DeepSeek encourages wider adoption, it may expand the revenue pool for these firms, consistent with the Jevons' paradox (Alcott 2005). As such, it remains an open question to be examined.

To test whether the existing GenAI providers and their hardware suppliers indeed suffered, I examine the market impact of DeepSeek on GenAI providers (i.e., Apple, Microsoft, Alphabet, Amazon, Meta, Nvidia, and Tesla) and their hardware suppliers (i.e., firms in Fama–French 48 "Computers" and "Electronic Equipment" industries). Panel A of Table 9 presents the summary statistics. The average CAR during the event window [0,5] is -0.046 and is statistically significant at 1% level. Panel B reports the correlation matrix and shows that *GenAI Exposure* is negatively correlated with *CAR [0,5]* at 1% level.

Panel C presents the regression results. In Columns (1) and (2), I use *CAR [0,5]* as the outcome variable. In Columns (3) and (4), I use *CAR [3,5]* (a three-day window that ends on January 28, 2025) as an alternative outcome variable. The coefficient on *GenAI Exposure* is negative and statistically significant in all four columns (t = -2.48, -1.70, -2.67, and -3.08, respectively), indicating that firms that discussed GenAI in the pre-event period experience a more negative market reaction. These findings from the GenAI supply side help explain that the adopter-side premium could be attributed to DeepSeek's disruptive entry.

*5.5 Evidence from China*

Because DeepSeek is a developed by a Chinese firm and is widely portrayed in the business press as a catalyst for GenAI ecosystem all around the world, it is therefore important to test whether the baseline valuation effect can be generalized to Chinese firms. I replicate the baseline analyses for Chinese firms using a text-based *GenAI Exposure* measure derived from annual reports in 2023.[38] Table 10 presents the results. The coefficient on *GenAI Exposure* remains positive and statistically significant in all three columns (t = 10.85, 10.53, and 4.99, respectively), suggesting that the main findings based on U.S. firms also generalize to Chinese firms. The results from a chow-test indicate stronger results for Chinese firms than for the U.S. sample (untabulated).

To further understand whether Chinese firms indeed use DeepSeek and if so, what kind of business activities they use DeepSeek for, I conduct a field experiment on the two major Chinese investor interactive platforms (IIPs) by soliciting responses from the sample firms to a question regarding firms' adoption of DeepSeek specifically.[39] The question I posted on the platforms is:

*Dear Board Secretary, has your company already deployed DeepSeek? If so, in which specific business areas is it being used? What cost-and-benefit factors did you consider when integrating DeepSeek? If the company has plans for future deployments, which concrete business functions will DeepSeek serve? Look forward to your reply. Thank you!*

Since all questions posted on the IIPs need to be approved by the stock exchanges, and mass-

[38] The 2024 annual reports of Chinese firms are typically released from February to April 2025, which might contaminate the GenAI exposure measure if I also include 2024 annual reports to construct the exposure measure.
[39] See Lee and Zhong (2022) for a detailed discussion of the investor interactive platforms in China.

posted questions with similar content (though to different firms) can hardly be approved, I successfully posted 925 questions through March 6 to July 12, 2025, with most of the questions posted in March (90.0%). Hence, the analysis is presented as early evidence on the DeepSeek deployment in Chinese firms. I receive 922 responses of the 925 questions and then GPT API to identify whether the firm's response mentions it has already deployed DeepSeek (*Deployed*), it is going to deploy DeepSeek (*Plan*), it has noted DeepSeek and is closely monitoring the development of DeepSeek (*Attention*), it has already deployed DeepSeek or plan to use DeepSeek for internal management (*Management*), it has already deployed DeepSeek or plan to use DeepSeek for products and services (*Product and Service*).

Panel A of Appendix C presents the descriptive statistics on these variables. I find that 23.2% of the firms say they have already deployed DeepSeek, with another 31.6% planning to do so and 71.6% at least keeping it on their radar. Among intended uses, internal management (34.1%) is mentioned a bit more often than product-oriented applications (29.5%). Panel B of Appendix C presents several examples to illustrate the use of DeepSeek by Chinese firms in the early 2025.[40]

## 6. Conclusions

In this paper, I exploit the release of DeepSeek as an exogenous shock to the availability of open-source GenAI models and investigate its valuation impact on corporate adopters. DeepSeek's open-source models exhibit comparable performance with the leading OpenAI's closed-source model, while its open-source nature significantly reduces implementation costs and privacy concerns associated with close-source GenAI adoption. Using a disclosure-based measure of firms' pre-event exposure to GenAI, I document that U.S. firms with high GenAI exposure earn an average cumulative abnormal return of 1.2% over the [0,5] event window surrounding the release

[40] I also re-run the analyses in Section 5.1 for the Chinese firms based on the IIP transcripts to capture firms' discussion about the (potential) adoption of and the benefits of DeepSeek. The results still hold (untabulated).

date of DeepSeek, relative to low-exposure firms in the same industry. The effect is stronger for firms facing tighter financial constraints and for firms with proprietary information concerns, shedding light on the benefits of open-source GenAI in reducing implementation costs and mitigating privacy concerns. Analyses of post-event conference call transcripts confirm that high-exposure firms subsequently are more likely to articulate DeepSeek adoption and discuss its benefits. High-exposure firms are also more likely to incorporate DeepSeek's open-source model into the code they host on GitHub. Financial analysts revise earnings forecasts upward and business press use a more positive tone for high-exposure firms. In contrast, GenAI model providers and their hardware suppliers suffer negative market returns, suggesting that DeepSeek's open-source entry reshuffles value from upstream GenAI vendors to downstream GenAI adopters. Analysis of the Chinese firms yields even larger valuation effects.

Although the evidence is suggestive, this study provides early evidence on the value of open-source GenAI models by showing that open-source alternatives can unlock additional value for (existing and potential) corporate adopters once adoption barriers fall. The results enrich both the growing work on GenAI adoption and the nascent literature on the value of open-source software by documenting GenAI adopters' valuation gains from GenAI producers releasing model weights and enabling local deployment. Future research can utilize actual adoption data to examine whether firms indeed deploy open-source GenAI models after the release of DeepSeek and to understand how firms integrate GenAI into business and operations.

## References


Acemoglu, D., & Restrepo. P. (2018). The Race Between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment. *American Economic Review*, 108(6), 1488-542

Acemoglu, D. (2025). The Simple Macroeconomics of AI. *Economic Policy*, 40(121), 13-58.

Alcott, B. (2005). Jevons' Paradox. *Ecological Economics*, 54(1), 9-21.

Awyong, A., Cheng, Q., Deng, T., & Wang, R. (2024). Digitalization, Accounting Jobs, and Financial Reporting Quality. ABFER Working paper.

Bai, J. J., & Jin, W. (2024). Cloud Adoption and Firm Performance: Evidence from Labor Demand. SSRN Working Paper.

Banerji, G., A. Fitch, & Langley, K. (2025). The Day DeepSeek Turned Tech and Wall Street Upside Down. *Wall Street Journal*. Available at https://www.wsj.com/finance/stocks/the-day-deepseek-turned-tech-and-wall-street-upside-down-f2a70b69.

Babina, T., Fedyk, A., He, A., & Hodson, J. (2024). Artificial Intelligence, Firm Growth, and Product Innovation. *Journal of Financial Economics*, 151, 103745.

Bertomeu, J., Lin, Y., Liu, Y., & Ni, Z. (2025a). The Impact of Generative AI on Information Processing: Evidence from the Ban of ChatGPT in Italy. *Journal of Accounting and Economics*, 101782.

Bertomeu, J., Lin, Y., Liu, Y., & Ni, Z. (2025b). What Are the Costs of Delayed Technology Adoption? Evidence from Italy's Ban of ChatGPT. SSRN Working Paper.

Blankespoor, E., deHaan, E., & Li, Q. (2025). Generative AI in Financial Reporting. SSRN Working Paper.

Boston Consulting Group (BCG). (2024). Where's the Value in AI? Available at https://www.bcg.com/publications/2024/wheres-value-in-ai.

Bousquette, I. (2025a). U.S. Businesses Already Love DeepSeek. *Wall Street Journal*. Available at https://www.wsj.com/articles/u-s-businesses-already-love-deepseek-1679cde2.

Bousquette, I. (2025b). No One Knows How to Price AI Tools. *Wall Street Journal*. Available at https://www.wsj.com/articles/no-one-knows-how-to-price-ai-tools-f346ea8a.

Bradshaw, M. T., Ma, C., Yost, B., & Zou, Y. (2025). Generative AI Use by Capital Market Information Intermediaries: Evidence from Seeking Alpha. SSRN Working Paper.

Breuer, M., & DeHaan, E. D. (2024). Using and Interpreting Fixed Effects Models. *Journal of Accounting Research*, 62(4), 1183-1226.

Brynjolfsson, E., Li, D., & Raymond, L. (2025). Generative AI at Work. *Quarterly Journal of Economics*, 140(2), 889-942.

Chen, W., Shi, T. T., Srinivasan, S., & Zakerinia, S. (2024). The Value of AI Innovations in Non-IT Firms. ABFER Working Paper.

Chen, W., & Srinivasan, S. (2024). Going digital: Implications for Firm Value and Performance. *Review of Accounting Studies*, 29(2), 1619-1665.

Cheng, S., Crowley, R., Lin, P., & Zhao, Y. (2025). Algorithm Access for All: Information Processing Democratization via GitHub. Working Paper.

Cheng, Q., Lin, P., & Zhao, Y. (2025). Does Generative AI Facilitate Investor Trading? Evidence from ChatGPT Outages. SSRN Working Paper.

Cisco. (2024). The Cisco 2024 Data Privacy Benchmark Study. Available at https://www.cisco.com/c/dam/en_us/about/doing_business/trust-center/docs/cisco-privacy-benchmark-study-2024.pdf.

Coleman, B., Fronk, K. & Valentine, K. (2025). Corporate Adoption of an Open Innovation

Strategy: Evidence from GitHub. SSRN Working Paper.
Conti, A., Peukert, C., & Roche, M. (2025). Beefing IT Up for Your Investor? Engagement with Open Source Communities, Innovation, and Startup Funding: Evidence from GitHub. *Organization Science*, forthcoming.
Dedman, E., & Lennox, C. (2009). Perceived Competition, Profitability and the Withholding of Information about Sales and the Cost of Sales. *Journal of Accounting and Economics*, 48(2-3), 210-230.
Dyer, T. A., Glaeser, S., Lang, M. H., & Sprecher, C. (2024). The Effect of Patent Disclosure Quality on Innovation. *Journal of Accounting and Economics*, 77(2-3), 101647.
Ecker, F., Li, X., Li, Y., & Wu, F. (2025). How Stock Market Participants Use Generative Artificial Intelligence: Evidence from User-Platform Interaction Data. SSRN Working Paper.
Eisfeldt, A. L., Schubert, G., & Zhang, M. B. (2024). Generative AI and Firm Values. SSRN Working Paper.
Eloundou, T., Manning, S., Mishkin, P., & Rock, D. (2024). GPTs are GPTs: Labor Market Impact Potential of LLMs. *Science*, 384(6702), 1306-1308.
Emery, L. P., Lim, C., & Ye, S. (2024). The Private Value of Open-Source Innovation. SSRN Working Paper.
Gibney, E. (2025). China's Cheap, Open AI model DeepSeek Thrills Scientists. *Nature*, 638(8049), 13-14.
Glaeser, S. (2018). The Effects of Proprietary Information on Corporate Disclosure and Transparency: Evidence from Trade Secrets. *Journal of Accounting and Economics*, 66(1), 163-193.
Greenstein, S., & Nagle, F. (2014). Digital Dark Matter and the Economic Contribution of Apache. *Research Policy*, 43(4), 623-631.
Gurman, M. (2023). Samsung Bans Staff's AI Use After Spotting ChatGPT Data Leak. *Bloomberg*. Available at https://www.bloomberg.com/news/articles/2023-05-02/samsung-bans-chatgpt-and-other-generative-ai-use-by-staff-after-leak.
Hoberg, G., & Maksimovic, V. (2015). Redefining Financial Constraints: A Text-based Analysis. *Review of Financial Studies*, 28(5), 1312-1352.
Hoberg, G., & Phillips, G. (2016). Text-based Network Industries and Endogenous Product Differentiation. *Journal of Political Economy*, 124(5), 1423-1465.
Hoffmann, M., Nagle, F., & Zhou, Y. (2024). The Value of Open Source Software. SSRN Working Paper.
Ishii, J., & Xuan, Y. (2014). Acquirer-target Social Ties and Merger Outcomes. *Journal of Financial Economics*, 112(3), 344-363.
Jia, N., Li, N., Xu, D., & Zhang, Y. (2025). When IR Meets AI: The Impact of Generative AI on Corporate Information Environments. SSRN Working Paper.
Jin, W., & Bai, J. J. (2022). Cloud Adoption and Firm Performance: Evidence from Labor Demand. SSRN Working Paper.
Kim, J., & Valentine, K. (2021). The Innovation Consequences of Mandatory Patent Disclosures. *Journal of Accounting and Economics*, 71(2-3), 101381.
Koh, P. S., & Reeb, D. M. (2015). Missing R&D. *Journal of Accounting and Economics*, 60(1), 73-94.
Korkmaz, G., Calderón, J. B. S., Kramer, B. L., Guci, L., & Robbins, C. A. (2024). From GitHub to GDP: A Framework for Measuring Open Source Software Innovation. *Research Policy*, 53(3), 104954.

Krause, D. (2025a). DeepSeek and FinTech: The Democratization of AI and Its Global Implications. SSRN Working Paper.

Krause, D. (2025b). DeepSeek's Potential Impact on the Magnificent 7: A Valuation Perspective. SSRN Working Paper.

Lee, C. M., & Zhong, Q. (2022). Shall we Talk? The Role of Interactive Investor Platforms in Corporate Communication. *Journal of Accounting and Economics*, 74(2-3), 101524.

Lem, K. W. (2024). Data Analytics Strategy and Internal Information Quality. *Contemporary Accounting Research*, 41(2), 1376-1410.

Lerner, J., & Schankerman, M. (2013) The Comingled Code Open Source and Economic Development. *The MIT Press*.

Lerner, J., & Tirole, J. (2002). Some Simple Economics of Open Source. *Journal of Industrial Economics*, 50(2), 197-234.

Lerner, J., & Tirole, J. (2005). The Economics of Technology Sharing: Open Source and Beyond. *Journal of Economic Perspectives*, 19(2), 99-120.

Li, D. (2011). Financial Constraints, R&D Investment, and Stock Returns. *Review of Financial Studies*, 24(9), 2974-3007.

Lin, B. (2025). How DeepSeek's AI Stacks Up Against OpenAI's Model. *Wall Street Journal*. Available at https://www.wsj.com/articles/how-deepseeks-ai-stacks-up-against-openais-model-e938c3d6.

Lin, L., Chin, J., & Huang, R. (2025). China Is Quickly Eroding America's Lead in the Global AI Race. *Wall Street Journal*. Available at https://www.wsj.com/tech/ai/artificial-intelligence-us-vs-china-03372176.

Linn, M., & Weagley, D. (2024). Uncovering Financial Constraints. *Journal of Financial and Quantitative Analysis*, 59(6), 2582-2617.

McKinsey & Company. (2025). The State of AI: How Organizations Are Rewiring to Capture Value. Available at https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai.

Merkley, K. J. (2014). Narrative Disclosure and Earnings Performance: Evidence from R&D Disclosures. *The Accounting Review*, 89(2), 725-757.

Nagle, F. (2018). Learning by Contributing: Gaining Competitive Advantage Through Contribution to Crowdsourced Public Goods. *Organization Science*, 29(4), 569-587.

Nagle, F. (2019). Open Source Software and Firm Productivity. *Management Science*, 65(3), 1191-1215.

Rai, S., & Sun, Y. (2025). China Floods the World With AI Models After DeepSeek Success. *Bloomberg*. Available at https://www.bloomberg.com/news/articles/2025-03-25/china-floods-the-world-with-ai-models-after-deepseek-s-success.

Reinicke, C. (2025). Nvidia's $589 Billion DeepSeek Rout Is Largest in Market History. *Bloomberg*. Available at https://www.bloomberg.com/news/articles/2025-01-27/asml-sinks-as-china-ai-startup-triggers-panic-in-tech-stocks.

Robbins, C., Korkmaz, G., Guci, L., Calderón, J. B. S., & Kramer, B. (2021). A First Look at Open-source Software Investment in the United States and in Other Countries, 2009-2019. https://iariw.org/wp-content/uploads/2021/11/robbins-paper.pdf.

Rosenbush, S. (2025). DeepSeek's Big Question: Where Does AI's True Value Reside? *Wall Street Journal*. Available at https://www.wsj.com/articles/deepseeks-big-question-where-does-ais-true-value-reside-35a36cd5.

Stam, W. (2009). When Does Community Participation Enhance the Performance of Open

Source Software Companies? *Research Policy*, 38(8), 1288-1299.
Tucker, J., Wang, J., & Zhao, R. (2025). Corporate Use of AI and Subsequent Product Market Performance. SSRN Working Paper.
Verrecchia, R. E., & Weber, J. (2006). Redacted Disclosure. *Journal of Accounting Research*, 44(4), 791-814.
Yang, W. (2023). How Can Open Source Technology Ecosystem Create Value? Evidence from Investors' Reactions to Firms' GitHub Code Releases. SSRN Working Paper.

**Appendix A Variable Definitions**

| Variable Name | Definition |
|---|---|
| **Variables used in main analyses and sensitivity tests** | |
| Dependent Variables | |
| *CAR [0,5]* | The cumulative abnormal return (CAR) based on the Fama-French three-factor model during day 0 to day 5, with day 0 being January 21, 2025 (DeepSeek's release date, January 20 was not a trading day), and with day 5 being January 28, when the market attention to DeepSeek reached the peak. |
| *CAR [6,15]* | The cumulative abnormal return (CAR) based on the Fama-French three-factor model during day 6 to day 15. |
| *CAR [6,25]* | The cumulative abnormal return (CAR) based on the Fama-French three-factor model during day 6 to day 25. |
| *BHAR [0,5]* | The buy-and-hold abnormal returns (BHAR) based on the Fama-French three-factor model during day 0 to day 5. |
| *CAR [0,5]_Market_Adj* | The cumulative abnormal return (CAR) based on the market model during day 0 to day 5. |
| *CAR [-1,6]* | The cumulative abnormal return (CAR) based on the Fama-French three-factor model during day -1 to day 6. |
| *CAR [3,5]* | The cumulative abnormal return (CAR) based on the Fama-French three-factor model during day 3 to day 5. |
| *CAR [-6,-1]* | The cumulative abnormal return (CAR) based on the Fama-French three-factor model during day -6 to day -1. |
| Independent Variable | |
| *GenAI Exposure* | An indicator variable that equals one if firm *i* discussed GenAI-related topics at least once during the conference calls in 2023 and 2024, and zero otherwise. |
| *GenAI Exposure_Quartile* | A quartile rank of the raw GenAI exposure measure, which equals 0 if no GenAI exposure, 1, 2, and 3 if the raw GenAI exposure measure falls in the bottom, middle and top tercile, respectively. |
| *GenAI Exposure_Log* | The logarithm of one plus the raw GenAI exposure. |
| *GenAI Exposure_Presentation* | An indicator variable that equals one if firm *i* discussed GenAI-related topics at least once during the presentation session of the conference calls in 2023 and 2024, and zero otherwise. |
| *GenAI Exposure_10-K* | An indicator variable that equals one if firm *i* discussed GenAI-related topics at least once in SEC 10-K filings in 2023 and 2024, and zero otherwise. |
| *GenAI Exposure_Pseudo* | An indicator variable that equals one if firm *i* discussed tariffs-related topics at least once during the conference calls in 2023 and 2024, and zero otherwise. |
| Control Variables | |
| *AI Exposure* | An indicator that equals one if the average count of AI-related keywords in firm *i*'s conference call transcripts in 2023 and 2024 is above the sample median but firm *i* did not mention Gen-AI related keywords, and zero otherwise. |
| *Size* | The natural logarithm of market value of common equity (in thousands). |
| *Age* | Logarithm of firm age. Firm age is calculated as the number of years since the firm first appeared in Compustat. |

| | |
|---|---|
| *Lev* | The ratio of total liabilities to total assets. |
| *ROA* | The ratio of net income to total assets. |
| *Loss* | An indicator variable that equals one if the income before extraordinary items is negative, and zero otherwise. |
| *BM* | The ratio of the book value of equity to the market value of equity. |
| *Tangibility* | Net property, plant, and equipment scaled by total assets. |
| *R&D* | R&D expenditures scaled by total operating expenses. |
| *Advertising* | Advertising expenditures scaled by total operating expenses. |
| **Variables used in cross-sectional analyses** | |
| *Size_Small* | An indicator variable that equals one if *Size* is below the sample median, and zero otherwise. |
| *Age_Low* | An indicator variable that equals one if *Age* is below the sample median, and zero otherwise. |
| *Fin_Constrain_High* | An indicator variable if the financial constraints measure developed by Linn and Weagley (2024) is above the sample median, and zero otherwise. |
| *Competition_High* | An indicator variable that equals one if the product similarity score developed by Hoberg and Phillips (2016) is above the sample median, and zero otherwise. |
| *R&D_High* | An indicator variable that equals one if the R&D expenses scaled by operating expenses is above the sample median, and zero otherwise. |
| *Service_Ind* | An indicator variable that equals one if firm *i* is in the business services or personal services industry as defined in Fama-French 48 industry classification, and zero otherwise. |
| **Variables used in additional analyses** | |
| *DeepSeek Adoption* | An indicator variable that equals one if firm *i* disclosed during the conference calls after January 27, 2025 that it has adopted or plans to adopt DeepSeek and open-source models developed by followers, and zero otherwise. |
| *DeepSeek Benefit* | An indicator variable that equals one if firm *i* discussed the benefits of DeepSeek and open-source models developed by followers during the conference calls after January 27, 2025, and zero otherwise. |
| *DeepSeek GitHub* | An indicator that equals one if firm *i* has embedded DeepSeek models in their algorithms shared publicly on GitHub, and zero otherwise. |
| *Analyst_Forecast_Chg* | Percentage changes of analyst forecast consensus from pre-event to post-event periods. |
| *Media_Sentiment_Chg* | Percentage changes of average sentiment of media reporting from pre-event to post-event periods. |

## Appendix B Examples of Generative AI Discussions in Conference Call Transcripts

1. **Gartner Inc. (FF48 Industry: Business Services)**

- *…we see generative AI as being really helpful for our business. As you said, there are a lot of internal efficiencies where we've had, 5 years ago, we had teams of humans combing through publicly available information. Now we actually, today, use generative AI to improve our efficiency on those kinds of things… so it will help our internal efficiencies.*
- *I'd say we're highly differentiated from kind of the public information you get because we have a lot of proprietary information, proprietary insights. We have a research process, which is quite important in generating these proprietary insights. …So we say generative AI has really been a lot of help both with internal efficiencies with probably a better interface with our clients, helping clients with it, et cetera.*

2. **Airbnb Inc. (FF48 Industry: Restaurants, Hotels, Motels)**

- *The base models, the large language models, think of those as GPT-4. Google has a couple of base models, Microsoft reaches Entropic… And so think of that as essentially like building a highway. It's a major infrastructure project. And we're not going to do that. We're not an infrastructure company. But we're going to build the cars on the highway. In other words, we're going to design the interface and the tuning of the model on top of AI, on top of the base model.*
- *We are now building AI into our product. GPT-4 is available in our app. So we're going to be building GPT-4 into our interface.*

3. **John Wiley & Sons Inc. (FF48 Industry: Printing and Publishing)**

- *…and in short, we believe that these technologies present far more opportunity than risk.*
- *The second area we focused on is product and publishing innovation. Today, we are actively building and deploying AI editing tools to improve speed and quality of our journal content and reduce unit costs through process automation in the value chain. To accomplish this, we have built an award-winning AI R&D team and are actively working with an international AI advisory team that includes professors, Ph.D. researchers and AI thought leaders.*
- *These efforts are paying off. For example, we recently launched an internal pilot of our AI-powered article matching engine to help authors get published faster and in the right journals.*

4. **Walmart Inc. (FF48 Industry: Retail)**

- *Here are some recent examples of us being tech-powered. Our new generative AI-powered search on the Walmart U.S. app, which rolled out to iOS users last month and is coming to Android users this month, is a great example. One of those popular searches this month was, "Help me buy a Valentine's Day gift." And rather than searching separately for things like chocolates, a card, jewelry, flowers, the search returns a list of results that are relevant and curated.*

5. **Thermo Fisher Scientific Inc. (FF48 Industry: Measuring and Control Equipment)**

- *Here are some recent examples of us being tech-powered. Our new generative AI-powered search on the Walmart U.S. app, which rolled out to iOS users last month and is coming to Android users this month, is a great example. One of those popular searches this month was, "Help me buy a Valentine's Day gift." And rather than searching separately for things like chocolates, a card, jewelry, flowers, the search returns a list of results that are relevant and curated.*

## Appendix C Firms' Responses to Investors' Questions on DeepSeek Adoption

This appendix provides descriptive statistics on the attributes of firms' responses to a question posted on their investor interactive platforms (Panel A) and examples of firms' responses (Panel B). The question I posted on the platforms is:

> Dear Board Secretary, has your company already deployed DeepSeek? If so, in which specific business areas is it being used? What cost-and-benefit factors did you consider when integrating DeepSeek? If the company has plans for future deployments, which concrete business functions will DeepSeek serve? We, as investors, look forward to your reply. Thank you!

The variables in Panel A are defined as follows: *Deployed* is an indicator variable that equals one if the firm's response mentions it has already deployed DeepSeek, and zero otherwise. *Plan to deploy* is an indicator variable that equals one if the firm's response mentions it is going to deploy DeepSeek, and zero otherwise. *Attention* is an indicator variable that equals one if the firm's response mentions it has noted DeepSeek and is closely monitoring the development of DeepSeek, and zero otherwise. *Management* is an indicator variable that equals one if the firm's response mentions it has already deployed DeepSeek or plan to use DeepSeek for internal management, and zero otherwise. *Product and Service* is an indicator variable that equals one if the firm's response mentions it has already deployed DeepSeek or plan to use DeepSeek for products and services, and zero otherwise.

*Panel A: Descriptive statistics on the attributes to firms' responses*

| | N | Mean | Std. Dev. | Q1 | Median | Q3 |
|---|---|---|---|---|---|---|
| *Deployed* | 925 | 0.232 | 0.423 | 0 | 0 | 0 |
| *Plan to deploy* | 925 | 0.316 | 0.465 | 0 | 0 | 1 |
| *Attention* | 925 | 0.716 | 0.451 | 0 | 1 | 1 |
| *Management* | 925 | 0.341 | 0.474 | 0 | 0 | 1 |
| *Product and Service* | 925 | 0.295 | 0.456 | 0 | 0 | 1 |

*Panel B: Examples of firms' responses*

**Response that indicting the firm has already deployed DeepSeek**

(Baoshan Iron & Steel Co., Ltd. [600019])

Thank you for your interest. Our parent company, China Baowu, has recently launched the "2526 Project" and has already completed an on-premises deployment of the full-capacity version of DeepSeek R1, which has been initially integrated with Baowu's steel-industry large model. As one of the pilot units, Baoshan Iron & Steel has taken the lead in trial applications of both the steel large model and a generative-AI assistant. Since launching our "AI+" strategy last year, we have worked with several leading large-model providers and have achieved early results in blast-furnace large models and hot-rolling production. Looking toward 2025, we will further deepen the use of large-model technology in core areas such as business decision-making, product quality, and production control, driving management transformation and industrial upgrading—and injecting intelligent momentum into the company's high-quality growth.

**Response that indicting the firm plan to deploy DeepSeek**

(Shenzhen Desay Battery Technology Co. Ltd. [000049])

Hello! Our company attaches great importance to the development and application of artificial-intelligence technologies and is going to implement the DeepSeek large model in internal management to boost operational efficiency and drive high-quality growth. Thank you!.

**Response that indicting the firm has noted DeepSeek**

(Shanxi Huayang New Material Co., Ltd. [600281])

Dear investor, thank you for your interest. The company has not yet deployed DeepSeek. We closely monitor emerging technologies in our industry—including those related to DeepSeek—and, when considering their adoption, we conduct a comprehensive assessment of business needs, cost-benefit performance, technological maturity, and data security. In the future, should rigorous research and analysis show that DeepSeek or similar technologies align with our strategic development goals and prove cost-effective, we will consider deploying them at an appropriate time and will define specific application areas based on our actual business requirements. Thank you for your continued attention and support.

**Response that indicting the firm use DeepSeek for internal management**
(Guangzhou Baiyun International Airport [600004])
The company closely follows emerging technology trends and has formed a team to study and begin developing DeepSeek applications, while evaluating costs and benefits and drafting an implementation plan. Current efforts are focused on internal management to improve managerial efficiency. In the future, the company will extend DeepSeek, in a planned and phased manner, to areas such as production operations and passenger services to fully leverage its technical advantages and enhance overall operational efficiency and service quality. Thank you.

**Response that indicting the firm use DeepSeek for products and services**
(Huangshan Tourism Development Co.,Ltd. [600054])
Hello! The company has been closely monitoring developments and opportunities in the tourism industry and actively leveraging digital-intelligence technologies to enhance management and upgrade services. Our subsidiary Anhui Tuma Technology has integrated the Huangshan AI Travel Assistant with the DeepSeek large model, upgrading intelligent interaction, scenario-based services, and companion travel experiences, and it will keep optimizing and deepening this application to continually improve the visitor experience. Thank you for your interest in the company.

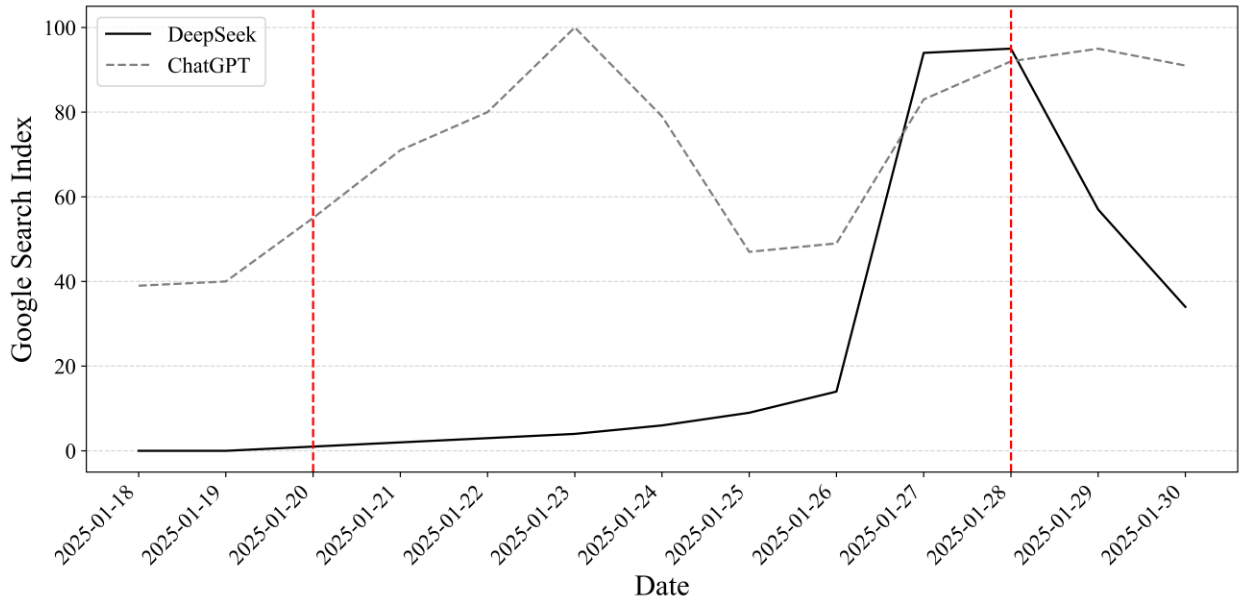


**Figure 1 Google Trends Interest Over Time: DeepSeek versus ChatGPT[41]**

[41] The Google Search Index represents relative search interest, where 100 is the peak popularity for the selected terms (i.e., "ChatGPT" and "DeepSeek") during the specified period (January 18-30, 2025, Eastern Time Zone). The Google Search Index for "DeepSeek" on January 18-19, 2025, was reported as "<1" by Google Trends. For plotting Figure 1, these values were replaced with 0 (See https://trends.google.com/trends/explore?date=2025-01-18%202025-01-30&geo=US&q=ChatGPT,DeepSeek&hl=en-US).

**Table 1**
**Sample Selection**

This table presents the sample selection process. The final sample for the cumulative abnormal return analyses includes 3,474 firm-level observations.

| | # Observations |
|---|---|
| Initial sample: Firm-level observations with Compustat financial data | 5,847 |
| Less: | |
| Observations with missing cumulative abnormal return data from CRSP | (845) |
| Observations with missing conference call transcripts data from Capital IQ during 2023 and 2024 | (1,211) |
| Observations with missing control variables | (98) |
| Generative AI providers and firms who are in computers and chips industries (according to Fama-French 48 industry code). | (219) |
| Final sample | 3,474 |

**Table 2**
**Descriptive Statistics and Correlation Matrix**

*Panel A: Descriptive statistics*

| Variable | N | Mean | SD | p25 | p50 | p75 |
|---|---|---|---|---|---|---|
| Variables used in the main event study analyses | | | | | | |
| *CAR [0,5]* | 3,474 | 0.002 | 0.077 | -0.029 | 0.005 | 0.038 |
| *GenAI Exposure* | 3,474 | 0.211 | 0.408 | 0.000 | 0.000 | 0.000 |
| *AI Exposure* | 3,474 | 0.288 | 0.453 | 0.000 | 0.000 | 1.000 |
| *Size* | 3,474 | 14.028 | 2.366 | 12.523 | 14.168 | 15.679 |
| *Age* | 3,474 | 2.879 | 0.818 | 2.197 | 2.944 | 3.497 |
| *Lev* | 3,474 | 0.622 | 0.319 | 0.406 | 0.597 | 0.806 |
| *ROA* | 3,474 | -0.076 | 0.306 | -0.077 | 0.012 | 0.057 |
| *Loss* | 3,474 | 0.390 | 0.488 | 0.000 | 0.000 | 1.000 |
| *BM* | 3,474 | 0.654 | 0.867 | 0.191 | 0.480 | 0.894 |
| *Tangibility* | 3,474 | 0.222 | 0.249 | 0.029 | 0.120 | 0.334 |
| *R&D* | 3,474 | 0.068 | 0.161 | 0.000 | 0.000 | 0.057 |
| *Advertising* | 3,474 | 0.013 | 0.036 | 0.000 | 0.000 | 0.007 |

## Table 2 (cont'd)

*Panel B: Correlation Matrix*

| | *CAR [0,5]* | *GenAI Exposure* | *AI Exposure* | *Size* | *Age* | *Lev* | *ROA* | *Loss* | *BM* | *Tangibility* | *R&D* |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *GenAI Exposure* | 0.121*** | | | | | | | | | | |
| *AI Exposure* | -0.032* | -0.329***[42] | | | | | | | | | |
| *Size* | 0.089*** | 0.161*** | 0.008 | | | | | | | | |
| *Age* | -0.002 | -0.091*** | 0.023 | 0.409*** | | | | | | | |
| *Lev* | 0.001 | -0.052*** | 0.038** | -0.046*** | 0.009 | | | | | | |
| *ROA* | 0.122*** | 0.087*** | 0.003 | 0.489*** | 0.330*** | -0.170*** | | | | | |
| *Loss* | -0.029* | -0.009 | 0.004 | -0.530*** | -0.388*** | 0.064*** | -0.571*** | | | | |
| *BM* | 0.010 | -0.076*** | -0.024 | -0.238*** | -0.011 | -0.327*** | 0.090*** | 0.056*** | | | |
| *Tangibility* | -0.155*** | -0.174*** | 0.003 | 0.058*** | 0.126*** | 0.012 | 0.101*** | -0.077*** | 0.021 | | |
| *R&D* | -0.014 | 0.000 | -0.057*** | -0.328*** | -0.242*** | 0.061*** | -0.638*** | 0.388*** | -0.159*** | -0.171*** | |
| *Advertising* | 0.068*** | 0.123*** | 0.047*** | -0.087*** | -0.112*** | 0.042** | 0.00500 | 0.064*** | -0.079*** | -0.050*** | 0.015 |

[42] The negative correlation between *GenAI Exposure* and *AI Exposure* is because *AI Exposure* takes a value of one only when (1) *GenAI Exposure* takes a value of zero and at the same time (2) the average count of (non-GenAI) AI-related keywords in firm *i*'s conference call transcripts in 2023 and 2024 is above the sample median.

**Table 2 (cont'd)**
**Industry distribution of GenAI exposure**

*Panel C: Industry Distribution of GenAI Exposure*[43]

| FF 48_Ind | Industry Name | # All firms | # GenAI Exposure | Ratio |
|---|---|---|---|---|
| 8 | Printing and Publishing | 10 | 8 | 80.00% |
| 34 | Business Services | 517 | 348 | 67.31% |
| 48 | Others | 71 | 27 | 38.03% |
| 33 | Personal Services | 53 | 17 | 32.08% |
| 37 | Measuring and Control Equipment | 60 | 18 | 30.00% |
| 32 | Communication | 69 | 19 | 27.54% |
| 42 | Retail | 150 | 37 | 24.67% |
| 43 | Restaurants, Hotels, Motels | 57 | 12 | 21.05% |
| 45 | Insurance | 116 | 21 | 18.10% |
| 40 | Transportation | 117 | 21 | 17.95% |
| 22 | Electrical Equipment | 63 | 11 | 17.46% |
| 11 | Healthcare | 58 | 10 | 17.24% |
| 47 | Trading | 265 | 41 | 15.47% |
| 46 | Real Estate | 33 | 5 | 15.15% |
| 23 | Automobiles and Trucks | 67 | 10 | 14.93% |
| 3 | Candy & Soda | 7 | 1 | 14.29% |
| 44 | Banking | 249 | 35 | 14.06% |
| 9 | Consumer Goods | 52 | 7 | 13.46% |
| 21 | Machinery | 98 | 13 | 13.27% |
| 41 | Wholesale | 91 | 11 | 12.09% |
| 31 | Utilities | 90 | 10 | 11.11% |
| 26 | Defense | 11 | 1 | 9.09% |
| 7 | Entertainment | 47 | 4 | 8.51% |
| 30 | Petroleum and Natural Gas | 114 | 8 | 7.02% |
| 19 | Steel Works Etc | 31 | 2 | 6.45% |
| 6 | Recreation | 19 | 1 | 5.26% |
| 27 | Precious Metals | 40 | 2 | 5.00% |
| 12 | Medical Equipment | 146 | 7 | 4.79% |
| 24 | Aircraft | 22 | 1 | 4.55% |
| 13 | Pharmaceutical Products | 381 | 17 | 4.46% |
| 28 | Non-Metallic and Industrial Metal Mining | 24 | 1 | 4.17% |
| 17 | Construction Materials | 52 | 2 | 3.85% |
| 10 | Apparel | 34 | 1 | 2.94% |
| 14 | Chemicals | 73 | 2 | 2.74% |
| 18 | Construction | 43 | 1 | 2.33% |
| 2 | Food Products | 49 | 1 | 2.04% |

[43] For 95 firms in Agriculture, Beer & Liquor, Tobacco Products, Rubber and Plastic Products, Textiles, Fabricated Products, Shipbuilding, Railroad Equipment, Coal, Business Supplies, Shipping Containers industries, they did not mention GenAI-related keywords during conference calls in 2023 and 2024.

**Table 3**
**Regressions of Abnormal Event Returns on GenAI Exposure**

This table reports the regression results of cumulative abnormal return (CAR) based on the following regression:

$$CAR\,[0,5]_i = \beta_0 + \beta_1 GenAI\ Exposure_i + Controls_i + \theta_{ind} + \varepsilon_i$$

*CAR [0,5]* is the cumulative abnormal return (CAR) based on the Fama-French three-factor model during day 0 to day 5. *GenAI Exposure* is an indicator variable that equals one if firm $i$ discussed GenAI-related topics at least once during the conference calls in 2023 and 2024, and zero otherwise. The sample includes 3,474 firm-level observations. See Appendix A for the variable definitions. Intercepts are included but not tabulated. The *t*-statistics in parentheses are based on robust standard errors. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively, based on two-sided tests.

| Dependent variable = | | *CAR [0,5]* | |
|---|---|---|---|
| | (1) | (2) | (3) |
| ***GenAI Exposure*** | **0.023*****| **0.011***** | **0.012***** |
| | **(6.80)** | **(3.02)** | **(2.63)** |
| *AI Exposure* | 0.001 | -0.002 | -0.002 |
| | (0.46) | (-0.64) | (-0.58) |
| *Size* | | 0.003*** | 0.003*** |
| | | (3.28) | (3.40) |
| *Age* | | -0.002 | -0.001 |
| | | (-1.19) | (-0.44) |
| *Lev* | | 0.011** | 0.006 |
| | | (2.31) | (1.13) |
| *ROA* | | 0.049*** | 0.047*** |
| | | (4.76) | (4.75) |
| *Loss* | | 0.010*** | 0.008** |
| | | (2.69) | (2.04) |
| *BM* | | 0.004** | 0.006** |
| | | (2.01) | (2.45) |
| *Tangibility* | | -0.045*** | -0.034*** |
| | | (-8.69) | (-3.70) |
| *R&D* | | 0.040*** | 0.017 |
| | | (2.72) | (1.05) |
| *Advertising* | | 0.118*** | 0.049 |
| | | (2.69) | (1.05) |
| FE | NO | NO | FF48 |
| N | 3,474 | 3,474 | 3,474 |
| Adj. $R^2$ | 0.014 | 0.059 | 0.124 |

## Table 4
## Regressions of Abnormal Event Returns on GenAI Exposure – Sensitivity Tests

*Panel A: Abnormal Event Returns during Longer Windows in the Following Weeks*

Panel A of Table 4 reports the sensitivity tests using abnormal event returns during longer windows in the following weeks. Column (1) reports the results using cumulative abnormal return based on the Fama-French three-factor model during day 6 to day 15 (*CAR [6,15]*) as an alternative measure of abnormal event returns. Column (2) reports the results using cumulative abnormal return based on the Fama-French three-factor model during day 6 to day 25 (*CAR [6,25]*) as an alternative measure of abnormal event returns. Controls and intercepts are included but not tabulated. The *t*-statistics in parentheses are based on robust standard errors. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively, based on two-sided tests.

| Dependent variable = | *CAR [6,15]* | *CAR [6,25]* |
|---|---|---|
| | (1) | (2) |
| *GenAI Exposure* | 0.011* | 0.022*** |
| | (1.96) | (2.76) |
| *Controls* | YES | YES |
| FE | FF48 | FF48 |
| N | 3,471 | 3,471 |
| Adj. $R^2$ | 0.047 | 0.046 |

*Panel B: Alternative Measures of Abnormal Event Returns*

Panel B of Table 4 reports the sensitivity tests using alternative measures of abnormal event returns. Column (1) reports the results using *BHAR [0,5]* (Buy-and-Hold Abnormal Returns) as an alternative measure of abnormal event returns based on the Fama-French three-factor model during day 0 to day 5. Column (2) reports the results using cumulative abnormal return based on market model as an alternative measure of abnormal event returns during day 0 to day 5. Controls and intercepts are included but not tabulated. The *t*-statistics in parentheses are based on robust standard errors. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively, based on two-sided tests.

| Dependent variable = | *BHAR [0,5]* | *CAR [0,5]_Market_Adj* |
|---|---|---|
| | (1) | (3) |
| *GenAI Exposure* | 0.011*** | 0.013*** |
| | (2.58) | (3.06) |
| *Controls* | YES | YES |
| FE | FF48 | FF48 |
| N | 3,474 | 3,474 |
| Adj. $R^2$ | 0.128 | 0.124 |

*Panel C: Alternative Event Windows*

Panel C of Table 4 reports the sensitivity tests using alternative event windows. Column (1) reports the results using cumulative abnormal return based on the Fama-French three-factor model during an extended event window from day -1 to day 6 (*CAR [-1,6]*). Column (2) reports the results using cumulative abnormal return based on the Fama-French three-factor model during an alternative event window from day 3 to day 5, with day 5 being Jan 28, 2025, when the market attention to DeepSeek reached the peak. Controls and intercepts are included but not tabulated. The *t*-statistics in parentheses are based on robust standard errors. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively, based on two-sided tests.

| Dependent variable = | *CAR [-1,6]* | *CAR [3,5]* |
|---|---|---|
| | (1) | (3) |
| *GenAI Exposure* | 0.012*** | 0.007** |
| | (2.69) | (2.22) |
| *Controls* | YES | YES |
| FE | FF48 | FF48 |
| N | 3,474 | 3,472 |
| Adj. $R^2$ | 0.115 | 0.144 |

*Panel D: Alternative Measures of GenAI Exposure*

Panel D of Table 4 reports the sensitivity tests using alternative measures of GenAI exposure. Column (1) reports the results using a quartile rank of raw GenAI exposure measure following Chen and Srinivasan (2024), which equals 0 if no GenAI exposure, 1, 2, and 3 if the raw GenAI exposure measure falls in the bottom, middle and top tercile, respectively. Column (2) reports the results using the logarithm of one plus the raw GenAI exposure as an alternative measure of GenAI exposure. Controls and intercepts are included but not tabulated. The *t*-statistics in parentheses are based on robust standard errors. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively, based on two-sided tests.

| Dependent variable = | *CAR [0,5]* | *CAR [0,5]* |
|---|---|---|
| *GenAI Exposure* | *GenAI Exposure_Quartile* | *GenAI Exposure_Log* |
| | (1) | (2) |
| *GenAI Exposure* | 0.024*** | 0.013*** |
| | (3.64) | (3.75) |
| *Controls* | YES | YES |
| FE | FF48 | FF48 |
| N | 3,474 | 3,474 |
| Adj. $R^2$ | 0.126 | 0.126 |

*Panel E: Alternative Sources of Discussions about GenAI*

Panel E of Table 4 reports the sensitivity tests using alternative sources to identify GenAI-related discussions. Column (1) reports the results using the presentation session of conference calls to construct *GenAI Exposure*. Column (2) reports the results using the Item 1 business section of SEC 10-K filings to construct *GenAI Exposure*. Controls and intercepts are included but not tabulated. The *t*-statistics in parentheses are based on robust standard errors. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively, based on two-sided tests.

| Dependent variable = | *CAR [0,5]* | *CAR [0,5]* |
|---|---|---|
| *GenAI Exposure* | *GenAI Exposure_Presentation* | *Using SEC 10-K filing* |
| | (1) | (2) |
| *GenAI Exposure* | 0.014*** | 0.019** |
| | (3.17) | (2.27) |
| *Controls* | YES | YES |
| FE | FF48 | FF48 |
| N | 3,474 | 2,172 |
| Adj. $R^2$ | 0.125 | 0.119 |

*Panel F: Placebo Tests*

Panel F of Table 4 reports the sensitivity tests using placebo measures for key variables. Column (1) reports the results using a placebo GenAI exposure measure constructed based on tariffs-related keywords. Column (2) reports the results using a pseudo-event window from day -6 to day -1. Controls and intercepts are included but not tabulated. The *t*-statistics in parentheses are based on robust standard errors. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively, based on two-sided tests.

| Dependent variable = | *CAR [0,5]* | *CAR [-6,-1]* |
|---|---|---|
| *GenAI Exposure* | *GenAI Exposure_Pseudo* | *GenAI Exposure* |
| | (1) | (2) |
| *GenAI Exposure* | 0.000 | -0.006 |
| | (0.14) | (-1.36) |
| *Controls* | YES | YES |
| FE | FF48 | FF48 |
| N | 3,474 | 3,474 |
| Adj. $R^2$ | 0.122 | 0.058 |

**Table 5**
**Cross-sectional Analyses**

This table reports the regression results of cross-sectional analyses based on the following regression:

$$CAR[0,5]_i = \beta_0 + \beta_1 GenAI\ Exposure_i \times CSVar_i + \beta_2 GenAI\ Exposure_i + \beta_3 CSVar_i + Controls_i + \theta_{ind} + \varepsilon_i$$

In Panel A, *Size_Small* is an indicator variable that equals one if *Size* is below the sample median, and zero otherwise. *Age_Low* is an indicator variable that equals one if *Age* is below the sample median, and zero otherwise. *Fin_Constrain_High* is an indicator variable if the financial constraints measure developed by Linn and Weagley (2024) is above the sample median, and zero otherwise. In Panel B, *Competition_High* is an indicator variable that equals one if the product similarity score developed by Hoberg and Phillips (2016) is above the sample median, and zero otherwise. *R&D_High* is an indicator variable that equals one if the R&D expenses scaled by operating expenses above the sample median, and zero otherwise. *Service_Ind* is an indicator variable that equals one if firm *i* is in the business services or personal services industry as defined in Fama-French 48 industry classification, and zero otherwise. See Appendix A for the variable definitions. Intercepts are included but not tabulated. The *t*-statistics in parentheses are based on robust standard errors. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively, based on two-sided tests.

*Panel A: Cross-sectional Analyses on Financial Constraints*

| Dependent variable = | *CAR [0,5]* | | |
|---|---|---|---|
| *CSVar =* | *Size_Small* | *Age_Low* | *Fin_Constrain_High* |
| | (1) | (2) | (3) |
| ***GenAI Exposure × CSVar*** | **0.013*** | **0.022***** | **0.025***** |
| | **(1.76)** | **(3.36)** | **(3.03)** |
| *GenAI Exposure* | 0.007 | 0.001 | -0.003 |
| | (1.61) | (0.13) | (-0.46) |
| *CSVar* | 0.007* | 0.003 | -0.004 |
| | (1.81) | (0.67) | (-1.00) |
| *Controls* | YES | YES | YES |
| FE | FF48 | FF48 | FF48 |
| N | 3,474 | 3,474 | 2,508 |
| Adj. $R^2$ | 0.126 | 0.127 | 0.122 |

*Panel B: Cross-sectional Analyses on Proprietary Information Concerns*

| Dependent variable = | *CAR [0,5]* | | |
|---|---|---|---|
| *CSVar* = | *Competition_High* | *R&D_High* | *Service_Ind*[44] |
| | (1) | (2) | (3) |
| ***GenAI Exposure* × *CSVar*** | **0.024***** | **0.020***** | **0.036***** |
| | **(3.38)** | **(2.81)** | **(3.99)** |
| *GenAI Exposure* | -0.002 | 0.001 | 0.002 |
| | (-0.32) | (0.22) | (0.43) |
| *CSVar* | -0.002 | 0.005 | |
| | (-0.44) | (1.06) | |
| *Controls* | YES | YES | YES |
| FE | FF48 | FF48 | FF48 |
| N | 2,953 | 3,474 | 3,474 |
| Adj. $R^2$ | 0.131 | 0.128 | 0.129 |

[44] The coefficient on the standalone *CSVar* is subsumed by the industry FE.

## Table 6
## Tests of DeepSeek-related Discussions in the Post-event Period

Table 6 reports the regression results of DeepSeek-related discussions based on the following regression:

$$DeepSeek\ Adoption_i(DeepSeek\ Benefit_i) = \beta_0 + \beta_1 GenAI\ Exposure_i + Controls_i + \theta_{ind} + \varepsilon_i$$

*DeepSeek Adoption* is an indicator variable that equals one if firm *i* disclosed during the conference calls after January 20, 2025 that it adopts or plans to adopt DeepSeek and open-source models developed by followers, and zero otherwise. *DeepSeek Benefit* is an indicator variable that equals one if firm *i* discussed the benefits of DeepSeek and open-source models developed by followers during the conference calls after January 20, 2025. See Appendix A for the variable definitions. Intercepts are included but not tabulated. The *t*-statistics in parentheses are based on robust standard errors. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively, based on two-sided tests.

| Dependent variable = | *DeepSeek Adoption* | *DeepSeek Benefit* |
|---|---|---|
| | (1) | (2) |
| ***GenAI Exposure*** | **0.115*** ** | **0.096*** ** |
| | **(8.16)** | **(7.41)** |
| *AI Exposure* | 0.011** | 0.014** |
| | (2.04) | (2.52) |
| *Size* | 0.005** | 0.003 |
| | (2.07) | (1.41) |
| *Age* | -0.016*** | -0.010** |
| | (-3.68) | (-2.56) |
| *Lev* | -0.006 | -0.015 |
| | (-0.42) | (-1.10) |
| *ROA* | 0.016 | 0.026** |
| | (1.02) | (2.05) |
| *Loss* | -0.003 | 0.005 |
| | (-0.30) | (0.51) |
| *BM* | 0.008 | 0.011* |
| | (1.56) | (1.87) |
| *Tangibility* | 0.033* | 0.029 |
| | (1.84) | (1.44) |
| *R&D* | 0.044 | 0.054** |
| | (1.49) | (2.00) |
| *Advertising* | -0.119 | -0.128 |
| | (-1.17) | (-1.31) |
| FE | FF48 | FF48 |
| N | 3,128 | 3,128 |
| Adj. $R^2$ | 0.099 | 0.079 |

**Table 7**
**Use of DeepSeek in Firms' GitHub Repositories**

Table 7 reports the regression results of firms' use of DeepSeek in their GitHub repositories based on the following regression:

$$DeepSeek\ GitHub_i = \beta_0 + \beta_1 GenAI\ Exposure_i + Controls_i + \theta_{ind} + \varepsilon_i$$

*DeepSeek GitHub* is an indicator variable that equals one if firm *i* has embedded DeepSeek models in their algorithms shared publicly on GitHub, and zero otherwise. See Appendix A for the variable definitions. Intercepts are included but not tabulated. The *t*-statistics in parentheses are based on robust standard errors. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively, based on two-sided tests.

| Dependent variable = | *DeepSeek GitHub* |
|---|---|
| | (1) |
| ***GenAI Exposure*** | **0.024*** ** |
| | **(4.68)** |
| *AI Exposure* | -0.008*** |
| | (-4.71) |
| *Size* | 0.004*** |
| | (2.78) |
| *Age* | -0.004** |
| | (-2.15) |
| *Lev* | 0.000 |
| | (0.05) |
| *ROA* | 0.018** |
| | (2.50) |
| *Loss* | 0.007 |
| | (0.96) |
| *BM* | -0.000 |
| | (-0.00) |
| *Tangibility* | -0.014** |
| | (-2.00) |
| *R&D* | 0.055*** |
| | (3.29) |
| *Advertising* | -0.102** |
| | (-2.33) |
| FE | FF48 |
| N | 2,783 |
| Adj. $R^2$ | 0.058 |

**Table 8**
**Tests of Changes in Expectation of Information Intermediaries**

Table 8 reports the regression results of expectation changes of information intermediaries based on the following regression:

$$Analyst\_Forecast\_Chg_i (Analyst\_Forecast\_Chg_i) = \beta_0 + \beta_1 GenAI\ Exposure_i + Controls_i + \theta_{ind} + \varepsilon_i$$

*Analyst_Forecast_Chg* is measured as percentage changes of analyst forecast consensus from pre-event to post-event periods. *Media_Sentiment_Chg* is measured as percentage changes of average sentiment of media reporting from pre-event to post-event periods. See Appendix A for the variable definitions. Intercepts are included but not tabulated. The *t*-statistics in parentheses are based on robust standard errors. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively, based on two-sided tests.

| Dependent variable = | *Analyst_Forecast_Chg* | *Media_Sentiment_Chg* |
|---|---|---|
| | (1) | (2) |
| ***GenAI Exposure*** | **0.072*** | **0.511*** |
| | **(1.78)** | **(1.78)** |
| *AI Exposure* | -0.008 | -0.032 |
| | (-0.26) | (-0.16) |
| *Size* | 0.020* | -0.132*** |
| | (1.87) | (-3.07) |
| *Age* | -0.043** | -0.069 |
| | (-2.03) | (-0.57) |
| *Lev* | -0.061 | 0.243 |
| | (-1.03) | (0.80) |
| *ROA* | -0.221** | 0.184 |
| | (-2.19) | (0.44) |
| *Loss* | -0.215*** | -0.107 |
| | (-3.76) | (-0.41) |
| *BM* | -0.019 | 0.001 |
| | (-0.49) | (0.01) |
| *Tangibility* | -0.068 | 0.436 |
| | (-0.77) | (0.95) |
| *R&D* | 0.253 | -0.268 |
| | (1.14) | (-0.27) |
| *Advertising* | -0.388 | 2.375 |
| | (-0.96) | (0.68) |
| FE | FF48 | FF48 |
| N | 1,765 | 3,325 |
| Adj. $R^2$ | 0.048 | 0.006 |

## Table 9
## Analyses on the Sample of Generative AI Providers and Their Hardware Suppliers

*Panel A: Descriptive Statistics*

| Variable | N | Mean | SD | p25 | p50 | p75 |
|---|---|---|---|---|---|---|
| *CAR [0,5]* | 217 | -0.046[45] | 0.060 | -0.077 | -0.038 | -0.004 |
| *GenAI Exposure* | 217 | 0.438 | 0.497 | 0.000 | 0.000 | 1.000 |
| *AI Exposure* | 217 | 0.134 | 0.341 | 0.000 | 0.000 | 0.000 |
| *Size* | 217 | 14.345 | 2.541 | 12.424 | 14.258 | 16.359 |
| *Age* | 217 | 3.143 | 0.718 | 2.833 | 3.332 | 3.584 |
| *Lev* | 217 | 0.490 | 0.301 | 0.285 | 0.457 | 0.630 |
| *ROA* | 217 | -0.071 | 0.323 | -0.100 | 0.017 | 0.072 |
| *Loss* | 217 | 0.442 | 0.498 | 0.000 | 0.000 | 1.000 |
| *BM* | 217 | 0.528 | 0.786 | 0.140 | 0.357 | 0.660 |
| *Tangibility* | 217 | 0.163 | 0.138 | 0.065 | 0.116 | 0.198 |
| *R&D* | 217 | 0.098 | 0.126 | 0.027 | 0.068 | 0.121 |
| *Advertising* | 217 | 0.003 | 0.009 | 0.000 | 0.000 | 0.002 |

[45] T-test of *CAR [0,5]* shows that the mean value is statistically significant at 1% level.

**Table 9 (cont'd)**

*Panel B: Correlation Matrix*

| | *CAR [0,5]* | *GenAI Exposure* | *AI Exposure* | *Size* | *Age* | *Lev* | *ROA* | *Loss* | *BM* | *Tangibility* | *R&D* |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *GenAI Exposure* | -0.182*** | | | | | | | | | | |
| *AI Exposure* | 0.082 | -0.347*** | | | | | | | | | |
| *Size* | 0.042 | 0.427*** | 0.031 | | | | | | | | |
| *Age* | 0.155** | 0.026 | 0.080 | 0.293*** | | | | | | | |
| *Lev* | -0.055 | 0.064 | -0.020 | 0.016 | 0.014 | | | | | | |
| *ROA* | 0.128* | 0.164** | 0.021 | 0.461*** | 0.426*** | -0.143** | | | | | |
| *Loss* | -0.116* | -0.206*** | 0.114* | -0.502*** | -0.302*** | 0.048 | -0.567*** | | | | |
| *BM* | 0.121* | -0.230*** | 0.061 | -0.288*** | 0.062 | -0.264*** | 0.012 | 0.176*** | | | |
| *Tangibility* | 0.053 | -0.041 | -0.018 | 0.074 | 0.016 | 0.013 | 0.050 | -0.013 | 0.250*** | | |
| *R&D* | -0.226*** | -0.009 | -0.063 | -0.245*** | -0.317*** | 0.078 | -0.466*** | 0.314*** | -0.147** | -0.154** | |
| *Advertising* | 0.125* | 0.027 | -0.080 | 0.035 | -0.018 | 0.015 | -0.050 | 0.006 | -0.114* | -0.116* | 0.006 |

## Table 9 (cont'd)

*Panel C: Regressions of Abnormal Event Returns on GenAI Exposure*

This table reports the regression results of cumulative abnormal return (CAR) based on the following regression:

$$CAR[0,5]_i(or\ CAR[3,5]_i) = \beta_0 + \beta_1 GenAI\ Adoption_i + Controls_i + \theta_{ind} + \varepsilon_i$$

*CAR [0,5]* is the cumulative abnormal return (CAR) based on the Fama-French three-factor model during day 0 to day 5. *CAR [3,5]* is the cumulative abnormal return (CAR) based on the Fama-French three-factor model during day 3 to day 5, with day 5 being Jan 28, 2025, when the market attention to DeepSeek reached the peak. *GenAI Exposure* is an indicator variable that equals one if firm *i* discussed GenAI-related topics at least once during the conference calls in 2023 and 2024, and zero otherwise. The sample includes 217 firm-level observations from GenAI providers (i.e., Apple, Microsoft, Alphabet, Amazon, Meta, Nvidia, and Tesla as in Krause [2025]) and firms that are in computers and chips industries (according to Fama-French 48 industry code). See Appendix A for the variable definitions. Intercepts are included but not tabulated. The *t*-statistics in parentheses are based on robust standard errors. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively, based on two-sided tests.

| Dependent variable = | *CAR [0,5]* | | *CAR [3,5]* | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| ***GenAI Exposure*** | **-0.031**** | **-0.028*** | **-0.022***** | **-0.027***** |
| | **(-2.48)** | **(-1.70)** | **(-2.67)** | **(-3.08)** |
| *AI Exposure* | | 0.004 | | 0.002 |
| | | (0.20) | | (0.13) |
| *Size* | | -0.007* | | 0.000 |
| | | (-1.67) | | (0.19) |
| *Age* | | 0.011 | | 0.005 |
| | | (0.97) | | (0.80) |
| *Lev* | | 0.011 | | -0.007 |
| | | (0.46) | | (-0.45) |
| *ROA* | | -0.033 | | -0.009 |
| | | (-0.69) | | (-0.43) |
| *Loss* | | -0.034** | | -0.014 |
| | | (-2.15) | | (-1.33) |
| *BM* | | -0.004 | | 0.007 |
| | | (-0.35) | | (1.05) |
| *Tangibility* | | -0.016 | | 0.014 |
| | | (-0.28) | | (0.44) |
| *R&D* | | -0.064 | | -0.072* |
| | | (-0.80) | | (-1.78) |
| *Advertising* | | 0.515 | | 0.655** |
| | | (1.03) | | (2.40) |
| FE | NO | FF48 | NO | FF48 |
| N | 217[46] | 217 | 217 | 217 |
| Adj. $R^2$ | 0.023 | 0.047 | 0.028 | 0.117 |

[46] The sample is slightly (2 observations) smaller than 219 reported in Table 1 because I drop singleton firm from the regression analyses, as suggested by Breuer and deHaan (2024).

**Table 10**
**Regressions of Abnormal Event Returns on GenAI Exposure – Chinese Sample**

This table reports the regression results of cumulative abnormal return (CAR) based on Chinese listed firms using the following regression:

$$CAR[0,5]_i = \beta_0 + \beta_1 GenAI\ Exposure_i + Controls_i + \theta_{ind} + \varepsilon_i$$

*CAR [0,5]* is the cumulative abnormal return (CAR) based on the Fama-French three-factor model during day 0 to day 5. *GenAI Exposure* is an indicator variable that equals one if firm $i$ discussed GenAI-related topics at least once in the 2023 annual report, and zero otherwise. The sample includes 4,600 firm-level observations. See Appendix A for the variable definitions. Intercepts are included but not tabulated. The $t$-statistics in parentheses are based on robust standard errors. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively, based on two-sided tests.

| Dependent variable = | *CAR [0,5]* | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| ***GenAI Exposure*** | **0.032*****  | **0.033***** | **0.021***** |
| | **(7.70)** | **(8.09)** | **(4.56)** |
| *AI Exposure* | | 0.007*** | 0.004 |
| | | (3.34) | (1.64) |
| *Size* | | -0.003*** | -0.003*** |
| | | (-3.14) | (-3.05) |
| *Age* | | -0.000 | 0.000 |
| | | (-0.24) | (0.25) |
| *Lev* | | -0.006 | -0.006 |
| | | (-1.18) | (-1.10) |
| *ROA* | | 0.066*** | 0.083*** |
| | | (3.34) | (4.09) |
| *Loss* | | -0.013*** | -0.012*** |
| | | (-4.13) | (-3.71) |
| *BM* | | 0.003** | 0.003** |
| | | (2.03) | (2.30) |
| *Tangibility* | | 0.018*** | 0.022*** |
| | | (3.39) | (3.26) |
| *R&D* | | 0.351*** | 0.267*** |
| | | (8.00) | (5.27) |
| *Advertising* | | -0.070*** | -0.045** |
| | | (-3.94) | (-2.23) |
| FE | NO | NO | Industry |
| N | 4,600 | 4,600 | 4,600 |
| Adj. $R^2$ | 0.020 | 0.071 | 0.096 |

# Online Appendix

## Table OA-1

## Determinant Analyses of GenAI Exposure

This table reports the determinant analyses of firms' exposure to GenAI based on the following regression:

$$GenAI\ Exposure_i = \beta_0 + Controls_i + \theta_{ind} + \varepsilon_i$$

*GenAI Exposure* is an indicator variable that equals one if firm *i* discussed GenAI-related topics at least during the conference calls in 2023 and 2024, and zero otherwise. *AI Discussion* is an indicator variable that equals one if firm *i* discussed AI-related topics during the conference calls in 2023 and 2024, and zero otherwise. The sample includes 3,474 firm-level observations. See Appendix A for the variable definitions. Intercepts are included but not tabulated. The *t*-statistics in parentheses are based on robust standard errors. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively, based on two-sided tests.

| Dependent variable = | *GenAI Exposure* | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| *AI Discussion* | 0.297*** | 0.246*** | 0.149*** |
| | (30.77) | (24.69) | (14.60) |
| *Size* | | 0.032*** | 0.033*** |
| | | (8.59) | (9.33) |
| *Age* | | -0.066*** | -0.025*** |
| | | (-7.69) | (-3.11) |
| *Lev* | | -0.059*** | -0.029 |
| | | (-2.75) | (-1.34) |
| *ROA* | | 0.132*** | 0.072*** |
| | | (4.75) | (2.88) |
| *Loss* | | 0.054*** | 0.040** |
| | | (3.25) | (2.57) |
| *BM* | | -0.006 | 0.003 |
| | | (-0.75) | (0.35) |
| *Tangibility* | | -0.200*** | -0.206*** |
| | | (-9.21) | (-5.97) |
| *R&D* | | 0.193*** | 0.286*** |
| | | (4.42) | (6.57) |
| *Advertising* | | 0.810*** | 0.802*** |
| | | (3.67) | (3.61) |
| FE | NO | NO | FF48 |
| N | 3,474 | 3,474 | 3,474 |
| Adj. $R^2$ | 0.112 | 0.175 | 0.338 |

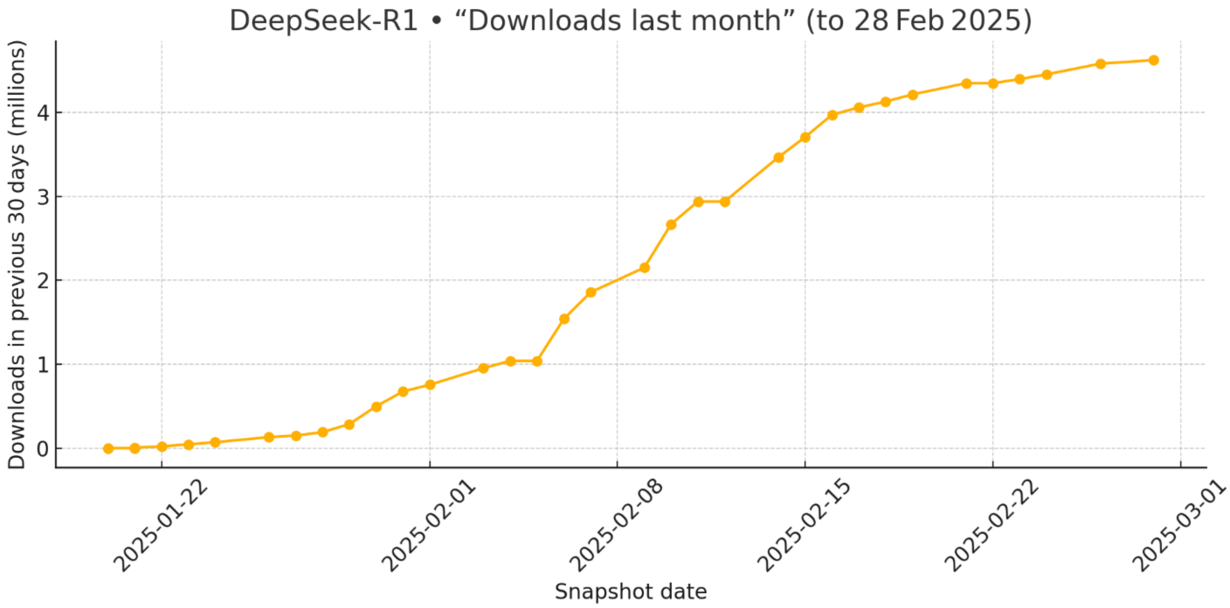


**Figure OA-1 Cumulative Downloads of DeepSeek-R1 model Over Past 30 days from Hugging Face Platform**[47]

[47] Source: I collect historical numbers of cumulative downloads of DeepSeek-R1 model over the past 30 days on the HuggingFace platform from WayBackMachine (https://web.archive.org/).