AI Technology Diffusion in the Stock Market

Runfeng Yang[a], Yufeng Mao[b,c], Massimiliano Caporin[d], Juan-Angel Jimenez-Martin[e]

Abstract

The AI technology diffusion over the past decade remains less understood. In this paper, we address this question by studying the information content of the "AI technology premium". We construct novel, unified measures of firm-level AI exposures through product and labor channels and analyze the US stock market from 2012 to 2024. We document a significant and highly time-varying AI technology risk premium. The product premium is positive and driven primarily by large firms, with positive shocks observed only after 2018. In contrast, the labor premium is weaker and is primarily driven by small firms. These patterns result from slow realization processes, varying investor attention, and the complex pricing of uncertainties related to AI adoption. Consequently, our results demonstrate that AI technology adoption represents a multifaceted process characterized by evolving investor perceptions and gradual realization of potential benefits, providing crucial insights for understanding market dynamics in the AI era.

Keywords: Artificial Intelligence, Asset Pricing, Technological Innovations, Return Predictability, Fama-MacBeth Regression

JEL Codes: G12, C51, G14

[a]Business School, Central South University, China.
[b]Center for Mathematics and Interdisciplinary Sciences, Fudan University, China.
[c]Shanghai Institute for Mathematics and Interdisciplinary Sciences, China.
[d]Department of Statistical Sciences, University of Padova, Italy.
[e]Instituto Complutense de Análisis Económico (ICAE), Universidad Complutense de Madrid, Spain.

1 Introduction

The rapid advancement of artificial intelligence (AI) constitutes one of the most significant technological shifts of the past decade. The diffusion of new technologies often correlates with specific stock price patterns (Pástor and Veronesi, 2009 and Greenwood and Jovanovic, 1999): initially, there is a bubble stage, which is typically followed by a decline accompanied by high volatility. These stock patterns evolve into a "technology risk premium" that encapsulates time-series and cross-sectional shocks arising from the adoption of the new technology

More precisely, the overall risk premium comprises two technology-related components. The first component involves exposure to a set of observed risk factors. For example, Babina et al. (2024b) show that firms with high AI investments exhibit higher market betas. The second component consists of a factor that carries unique technological information, which cannot be captured by the existing set of observed risk factors. Firms that adopt the new technology are inherently exposed to this novel risk factor, leading to subsequent stock price movements.

Consequently, analyzing the information content of the corresponding technology risk premium offers a method to track the technology diffusion process. One aspect distinguishing the technology risk premium from other risk premiums is the multifaceted and unobservable nature of technology adoption. Specifically, the impact of AI on firm operations can be categorized into two distinct aspects (Babina et al., 2024a and Eisfeldt et al., 2025): the product impact, which involves integrating AI into goods and services, and the labor impact, which entails using AI to enhance labor efficiency. Additionally, uncertainties may arise in the application of AI technology. However, much of the literature tends to focus on a single facet of AI adoption, neglecting its combined effects and the ensuing dynamic adoption process.[1] Consequently, several questions remain insufficiently understood: Do the labor and product impacts diffuse differently? Are investors recognizing the risks associated with AI adoption? How has AI technology adoption evolved? Addressing these questions is crucial because AI investment has increased tenfold over the past decade, suggesting a potential bubble stage.[2] Furthermore, AI technology represents a disruptive force, leading to system-wide transformations (Agrawal et al., 2024).

In this paper, we propose and analyze the "AI technology risk premium" as a novel framework to track the diffusion of AI technology. To address the complex nature of technology adoption, we construct unified measures that separately capture the impacts on both products and labor. These measures enable us to jointly describe the technological impact. The term "unified" refers to the fact that both measures are derived from the

[1] A growing body of literature has begun to explore AI's potential impact on firm valuation (e.g., Bertomeu et al., 2025, Eisfeldt et al., 2025 and Fotheringham and Wiles, 2023) or the overall economic productivity (e.g., Acemoglu, 2024, Korinek and Suh, 2024 and Brynjolfsson et al., 2021).

[2] https://hai.stanford.edu/ai-index/2025-ai-index-report/economy

same underlying definition of AI and constructed within a common framework, allowing for a comparative analysis of the relative importance of the product and labor channels.

For AI product exposure, we assess how the product space, represented by business descriptions in 10-K filings, is covered by the AI patent universe. Conversely, for AI labor exposure, we evaluate how a task is represented by this AI patent universe. We then aggregate this representation from the task level to the occupation level and further to the 3–6 digit NAICS code level based on the Bureau of Labor Statistics (BLS) survey data, which is subsequently extrapolated to the firm level. The AI patent universe, classified by Giczy et al. (2021), who employ machine learning and manual verification to identify AI-related patents, serves as a proxy for unobservable AI technology. This approach is superior to the bag-of-words method prevalent in literature, which arbitrarily selects a list of AI-related keywords. Our method objectively captures the time-varying nature of AI technology development.

Utilizing this unified AI measurement framework, we then examine the AI technology risk premium in the US market. Given that the rise of AI technology is a recent phenomenon – less than 1% of firms in the Russell 3000 mentioned AI in earnings calls in 2016, compared to 16% in 2023 – our sample period extends from January 2012 to December 2024.[3]

We first document several stylized facts regarding AI labor and product exposure. The information technology sector demonstrates the highest product exposure, whereas non-tech industries like healthcare and education also show significant exposure, reflecting AI technology's commercial applications. Conversely, agriculture and mining have the lowest product exposure. In terms of labor exposure, industries such as professional services exhibit high exposure, while retail and accommodation display low exposure. Additionally, at the occupational level, labor exposure is highest in roles involving data processing (e.g., computer science, social science) or the automation of low-skill tasks (e.g., administration, maintenance). In contrast, it remains low in jobs that require personal interactions (e.g., food service, personal care).

We identify a dynamic risk premium associated with AI technology exposure. A significant and time-varying positive premium is linked to AI product exposure. This premium was negligible in the early 2010s but experienced strong positive shocks after 2018, notable negative corrections post-2021, and further strong shocks after 2024. Moreover, the product premium is largely observed in value-weighted portfolios, suggesting that larger firms capture much of the AI premium, likely due to their greater implementation capacity. This positive product premium persists even when excluding technology-related companies, including the "Magnificent Seven" that have recently driven market index performance. In contrast, the premium related to AI labor exposure is generally weaker

[3]https://www.goldmansachs.com/insights/articles/ai-investment-forecast-to-approach-200-billion-globally-by-2025

and more complex. Before 2020, there was a modest positive AI labor premium, but this was offset by a significant negative labor premium during the COVID period. Recently, we have begun to observe a positive premium.

We explore the informational content and underlying mechanism of the observed AI product and labor premium. Our findings indicate a gradual realization of the AI product premium. Return predictability increases as the time horizon is extended from one month to 36 months, which aligns with recent discussions on the benefits of AI for companies. However, when examining the full-sample, we observe that positive predictability at the 3- to 6-month horizon is less pronounced than during the pre-COVID period. This suggests that the observed AI premium may not solely reflect a "company fundamental impact" but also involves behavioral factors that contributed to potential negative shocks during the COVID period.

To determine if the observed AI premium includes behavioral factors, we use the Google Trends Index as a proxy for AI-related attention and analyze its interaction with AI product and labor exposures. The interaction term between AI attention and AI product exposure is positive, whereas the interaction with AI labor exposure is negative. During periods of heightened AI attention, there is an associated increase in the product premium and a decrease in the AI labor premium. Additionally, we demonstrate that institutional investors dynamically adjust their holdings based on AI exposure. They tend to buy into the AI narrative during upward trends and sell during corrections, particularly around the COVID period.

We propose two explanations for the negative labor premium observed during the COVID period. First, we demonstrate that increased AI labor exposure correlates with negative changes in institutional holdings, partially accounting for the underperformance of firms with high AI labor exposure. Second, our analysis of the interaction between AI labor exposure and firm size during the COVID period reveals a negative coefficient for the interaction term. This finding suggests that smaller firms with significant AI labor exposure were the main contributors to the negative labor premium.

Our findings on the AI labor premium partially diverge from previous studies that report a positive valuation impact from AI-related labor effects (Chen and Wang, 2024), especially for generative AI technologies like ChatGPT (Eisfeldt et al., 2025). We reconcile our results with existing literature by distinguishing between the supporting and core roles of tasks when estimating AI labor exposure. During the pre-COVID period, the premium on supporting tasks was predominant. However, in recent years, the premium from core tasks has begun to have a positive influence, aligning with Eisfeldt et al. (2025)'s findings on generative AI technologies. This shift reflects how AI technology has evolved from being a supportive tool to a central driver across various industries.

The product impact reflects the application of AI technologies in developing new products or improving existing ones. In contrast, the labor impact pertains to changes in productivity of both labor and capital due to AI integration. However, these measures fail to consider the risks and uncertainties associated with AI adoption via the product channel. For instance, the risks of adopting AI (such as regulatory uncertainty and product failure) and risks of not adopting it (such as falling behind competitors). We refer to these risks collectively as "AI risk". We construct AI risk exposure by extracting the Risk Factors section (Item 1A) from 10-K filings and measuring its overlap with the AI patent universe. Our results indicate that AI risk is positively priced and does not subsume the product premium. AI risk has interacted positively with AI product exposure in the recent 2 years, but not in earlier periods. This suggests that although investors account for the unique regulatory and operational risks associated with AI, the AI product premium is not associated with AI risk, at least in earlier years.

Overall, our empirical findings, with a focus on AI technology, provide important evidence contributing to early discussions on the impact of technology adoption on stock prices (Pástor and Veronesi, 2009, Laitner and Stolyarov, 2003 and Greenwood and Jovanovic, 1999). Technology adoption typically unfolds in three stages. Initially, in the early stage, stock prices experience a bubble-like surge but tend to decline due to uncertainties surrounding the new technology. The significant positive shocks in AI product premiums post-2018, followed by corrections after 2021, align with model predictions.[4]

In the second stage, stock prices increase as the technology begins to materialize, correlating with the recent positive premiums observed for AI products and labor. The third stage is characterized by a decline in stock prices when increases in the discount rate exceed gains in cash flow. This stage was not captured in our sample since AI adoption is still in its nascent phases. Nonetheless, the situation may be more complex than model predictions suggest; strong positive shocks in the past 2 years hint at potential new rounds of bubbles. Additionally, we identify a market-behavior channel, represented by shocks associated with AI attention, which previous equilibrium models do not capture. However, during our sample period, the labor channel for AI technology realization remains limited.

We also investigate whether AI product premiums vary across different AI technologies. Following Giczy et al. (2021), we classify AI patents into eight categories and investigate whether these technologies produce heterogeneous product impacts. We construct technology-specific AI product exposure by combining patent classification data with our business-description-based AI exposure. Long-short portfolios based on different technologies reveal significant heterogeneity: computer vision, speech, and natural language processing (NLP) generate positive returns, whereas machine learning and planning/control show insignificant effects, with the former delivering nearly twice the average

[4]Though we explain, in the main analysis, that the negative correction involves investor's sentiment, another potential explanation for the negative correction is from increased concern for the uncertainty of AI technology adoption.

return of the latter.

We check whether our results are influenced by measurement errors in the AI exposure metrics. For AI product exposure, we consider two alternative measures: AI employment data fromBabina et al. (2024a), which captures AI through the share of AI-related employees, and the proportion of AI-related patents to total patents. For AI labor exposure, we employ text-embedding techniques to measure task-level exposure. Our conclusions remain robust across these alternative measures. Finally, we find limited evidence that our AI premium is driven by digitization or other firm-level risks identified by Hassan et al. (2019).

Related Literature. We contribute to the growing literature on AI technology adoption and firm value. Most studies find a positive relationship between the two (e.g., Bertomeu et al., 2025, Eisfeldt et al., 2025, Fotheringham and Wiles, 2023, Fotheringham and Wiles, 2023, and Rock, 2019). However, there is not yet a consensus (and sometimes contradictory views) on the possible channels through which AI adoption is realized, probably due to the varied empirical settings in dierent studies. Some studies relate the AI premium to the product channel. Babina et al. (2024a) show that production innovation (represented as the number of product patents and trademarks), instead of productivity improvement, is the main driver of firm growth from the commercial AI adoption. Similarly, Bertomeu et al. (2025) find that the banning of ChatGPT would indicate creative destruction for new firms that rely on the technology for business. Rock (2019) shows that AI-intensive companies rapidly gained market value following the launch of TensorFlow (an open-source deep learning package). Lui et al. (2022), using data from 2015 to 2019, show that the price of the stock falls by 1.77% on the day of the announcement of AI investments. They find that investors interpret AI investment announcements as unwelcome news for most firms.

Other studies associate the positive premium with the labor channel. Eisfeldt et al. (2025) find that the release of ChatGPT corresponds to a potential increase in firm productivity and a decrease in labor demand. According to Chen and Srinivasan (2023) the adoption of digital technologies, including machine learning and automation, weakly increases interim productivity. They also observe declines in sales growth conditional on digital activities.

Our study distinguishes itself from previous analyses in several key ways. First, we develop a unified empirical framework that simultaneously captures both the product and labor impacts of AI. Unlike prior studies relying on proprietary data, such as firm-level employment records, our measures are constructed entirely from publicly available information accessible to investors. Second, we extend beyond static snapshots to demonstrate that the market's pricing of AI has evolved through different phases. Initially, there was a period of negligible premiums, followed by increased optimism and positive shocks, and later episodes of correction in investor expectations. This time-series perspective, which traces how pricing dynamics unfold over time, is largely absent from existing literature. Third, we explicitly consider the role of behavioral factors and find that they play distinct roles in the product and labor channels. Additionally, to capture the risks and uncertainties associated with AI adoption through these channels, we construct and price a novel measure of "AI risk", showing that investors require compensation for these AI-related uncertainties. By integrating these behavioral and risk-based explanations, we offer a more holistic understanding of the forces driving AI technology impact in the capital market.

We relate our paper to theoretical studies on technology innovations and asset prices(e.g., Lin et al., 2020, Kogan and Papanikolaou, 2014, Papanikolaou, 2011, and Pástor and Veronesi, 2009), as well as to empirical studies on the asset pricing implications of technology innovation (e.g., Lee et al., 2019, Zhang, 2019, Hirshleifer et al., 2013, and Hsu, 2009). Our investigation into AI adoption and capital markets serves as a field experiment in technology diffusion. We find that the stage of technology adoption significantly influences how capital markets value it. Additionally, the slow diffusion of new technology delays the realization of its valuation premium.

The paper is organized as follows. In Section 2, we present a conceptual discussion on the relationship between AI exposure and stock performance. Section 3 covers the measurement of AI adoption levels and the data utilized for analysis. In Section 4, we provide the results from the portfolio sorts and the Fama-MacBeth regressions. We then explore potential driving forces behind the AI premium in Section 5. Section 6 outlines various robustness tests, including results from different measures of AI exposure and the impact of technology. Finally, we conclude our findings in Section 7.

2 AI Technology Diffusion and Stock Performance

AI is defined as "systems that display intelligent behavior by analyzing their environment and taking actions – with some degree of autonomy – to achieve specific goals" (Sheikh et al., 2023). The diffusion of AI technology since 2010 can be divided into two stages. In the first stage, the technology was predominantly used for commercial applications, including chatbots, industrial robots, recommendation algorithms, and predictive analytics. In the second stage, marked by the advent of large language models and generative AI, the technology became directly accessible to end users. Overall, the impact of AI technology can be categorized into two types: product impact and labor impact.

Product impact refers to the integration of AI technology for developing new products or enhancing existing ones. For instance, social media platforms employ AI algorithms to personalize content tailored to users. In the healthcare sector, AI assists in drug discovery and the prediction of chemical reactions. In the financial sector, AI can power personalized chatbots and enhance fraud detection. In the manufacturing sector, AI is utilized for quality control and supply chain management. Babina et al. (2024a) have shown that companies with higher AI investments are linked to a greater number of product patents and trademarks

The labor impact encompasses changes in the productivity of both labor and capital. AI contributes by automating various tasks; for instance, generative AI can undertake writing and summarization tasks. Noy and Zhang (2023) demonstrate that GPT enhances productivity in writing tasks. Historically, automation predominantly occurred in manufacturing, often through the use of industrial robots, such as those for automated inventory and quality control. This topic closely relates to the labor displacement literature (e.g., Eloundou et al., 2024; Acemoglu et al., 2022b; Acemoglu and Restrepo, 2020).

Regarding the potential mechanism, no existing equilibrium model adequately describes the relationship between AI technology adoption and stock prices. Most theoretical and empirical studies on technological innovation and stock valuation suggest that new technologies ultimately lead to improvements in firm productivity, which foster firm growth (e.g., Chen and Srinivasan, 2023, Zhang, 2019, Kogan and Papanikolaou, 2014, and Papanikolaou, 2011) or alter the risk profile (e.g., Lin et al., 2020). In theory, the impacts on labor and products can manifest through various channels that influence stock prices. For instance, when a company decides to adopt AI technology, investors reassess the firm's value through two primary channels: the cash flow channel and the discount rate channel.

In the cash flow channel, changes occur in expected future cash flows. For example, Babina et al. (2023) illustrate that AI investment can drive corporate innovation, leading to increased sales and firm growth. Earlier studies also document that AI fosters product innovation (e.g., Cockburn et al., 2018). Furthermore, Eisfeldt et al. (2025) demonstrate that the release of generative AI enhances labor efficiency, with the effect of ChatGPT on firm value primarily influencing expected future cash flows. Similarly, Chen and Wang (2024) find that AI innovation improves firms' demand for skilled labor, thereby contributing to firm growth. AI may also stimulate growth by enhancing production efficiency; this includes improvements such as better supply chain management.

In the discount rate channel, exposure to AI technologies alters a firm's risk profile. On one hand, AI adoption can introduce cash flow uncertainty, as the implementation of new technologies may fail. As indicated by Babina et al. (2023), higher AI investment is associated with lower systematic risk, suggesting that AI provides firms with valuable growth opportunities. Conversely, firms with greater AI exposure might be perceived as less risky, owing to a reduced likelihood of becoming technological laggards. However, firms heavily engaged in AI development might encounter heightened regulatory risks related to AI governance or increased product failure risks due to the immature application of technology. Additionally, behavioral factors could influence this scenario. Investors may exhibit AI-related attention or preferences, particularly given the recent surge of interest in generative AI.

All these factors may operate simultaneously, generating confounding or even opposing effects. Consequently, the relationship between AI and stock prices can evolve. While it is important to understand the equilibrium state, it is equally valuable to examine the process through which equilibrium is reached and consider when such a state is relevant. This consideration is particularly pertinent within the context of technological adoption, which often unfolds over several years. During this diffusion period, stock prices are subject to dynamics (e.g., Gârleanu et al., 2012; Pástor and Veronesi, 2009). However, comprehensive research investigating the overall impact of AI technologies on stock performance and their dynamic interactions remains lacking. Additionally, the risks associated with AI adoption are not yet addressed.

Building on the preceding discussion, our empirical design aims to address two key questions that remain unexplored in the existing literature. First, is the relationship between AI adoption and stock performance consistently positive over time, as commonly suggested by prior studies? Second, what factors drive this observed positive relationship? Specifically, have all the channels discussed above contributed to this relationship over the past decade?

3 Data and Measurement

3.1 Measuring the AI Exposure

A company's level of AI exposure represents an intangible asset, the quantification of which depends on the types of AI technologies utilized and the firm's stage of development. Furthermore, firm-level disclosures regarding AI adoption are often limited, especially during the early phases of technological diffusion. Together, these factors make measuring a firm's AI exposure inherently challenging.

Many studies capture product exposure through text mining of 10-K reports (e.g., Chen and Srinivasan, 2023, El Moujahid et al., 2023 and Mishra et al., 2022) and company announcements (Lui et al., 2022). Alternatively, some studies use employment-based data. For instance, Babina et al. (2024a) quantify a firm's AI investment by examining the proportion of employees with AI-related skills. Finally, several studies employ patent or R&D information (e.g., Chen and Wang, 2024, Kelly et al., 2021, Lin et al., 2020, Lee et al., 2019 and Hsu, 2009) or survey-based data (Acemoglu et al., 2022a) to represent technology adoption.

Other studies concentrate on assessing the labor impact of AI technology. The fundamental step involves evaluating whether a task can be replaced or influenced by AI. Subsequently, task-level exposure is aggregated to the industry level following this pathway:

AI → task → occupation → firm/industry. Earlier research employed crowd-sourcing services such as Amazon Mechanical Turk and CrowdFlower to recruit individuals to assess whether AI could affect a given task (e.g., Acemoglu et al. (2022b), Felten et al., 2019, Felten et al., 2018, and Brynjolfsson et al., 2018). Recently, studies have begun utilizing large language models (LLMs) to determine task-level exposure to AI technologies, particularly to assess whether a task can be replaced by generative AI (e.g., Eloundou et al., 2024 and Eisfeldt et al., 2025).

The selection of an appropriate AI indicator hinges on how the impact of AI is conceptualized. As discussed in the previous section, our objective is to capture the comprehensive impacts of AI technology. Relying on a single indicator is insufficient for accurately capturing the AI technology risk premium. To address this issue, we propose a unified framework that integrates both dimensions, as illustrated in 1. Detailed descriptions of the estimation procedures for AI labor and product exposure measures, along with illustrative examples and validation exercises, are provided in B and C. We begin by defining and measuring the unobservable development of AI technology.

We measure AI technological progress annually by examining AI-related patent grants from the previous decade, defining these as the AI patent universe for that year. To separate patents into AI and non-AI categories, we adopt the method proposed by Giczy et al. (2021) which utilizes machine learning techniques to classify patents into eight technology classes based on their texts, offering classification data extending back to 1976.[5] For comparison, Chen and Wang (2024) employ ChatGPT to categorize patents into four groups using their abstracts, while Webb (2020) adopts a keyword-search approach followed by manual verification. The underlying premise is that the features of AI technology represented in the AI patent universe provide a more objective categorization than a pre-defined, arbitrary bag-of-words.

Next, we construct the AI product measure as detailed in Appendix B. For each year m, we compute a $p_m \times 1$ firm-specific embedding $V_{i,m}$ that captures the semantic direction of firm i's business and a $p_m \times 1$ embedding $V_{\text{Patent},m}$ that captures the annual semantic direction of AI technology within the previously defined AI universe. The vectors are constructed from firms' 10-K business descriptions and the abstracts of AI patents, and both are ℓ_2-normalized. We define the AI product measure for firm i in year m as the cosine similarity:

$$\text{AI}_{i,m}^{\text{Prod}} = V_{i,m}^{\top} V_{\text{Patent},m}. \tag{1}$$

A higher cosine similarity value indicates greater AI product exposure. Firms may mention AI-related words in their 10-K reports for several reasons: (1) AI technologies or services constitute their core products (e.g., OpenAI, which develops advanced AI mod-

[5] Classification data are obtained from the U.S. Patent and Trademark Office: https://www.uspto.gov/ip-policy/economic-research/research-datasets/artificial-intelligence-patent-dataset.

els); (2) the firm integrates AI into its business model (e.g., online platforms employing AI-based recommendation systems); (3) AI enhances internal operations or governance (e.g., banks applying AI and machine learning algorithms for risk assessment, or manufacturers using AI for predictive maintenance); or (4) AI has become a key feature of their products (e.g., autonomous driving technologies). Consequently, the 10-K–based measure provides a reasonable proxy for a firm's product exposure to AI technologies.

The primary distinction between our data-driven method and a conventional bag-of-words approach is that the AI technology vector, $V_{\text{Patent},m}$, is composed of the semantic features of the AI patents universe. As illustrated in Table G.9, the most frequent nouns are "information", "computer", and "data", and similar terms. They represent the building blocks of AI systems rather than the label "AI" itself. One might contend that this method could also capture exposure to broader digital technologies. However, AI technology itself, as defined at the beginning of Section 2, is deeply associated with data processing and computing. There is a trade-off between accuracy and scope with this approach: using nouns like "data" might inadvertently capture other digital technologies, whereas limiting the measure to a few specific keywords could result in an AI product measure of zero for most companies. We address this issue by studying the effects of digitalization in Section 5 and by comparing alternative AI measures in Section 6.

To measure AI labor exposure, we adopt the task-occupation framework proposed by Eisfeldt et al. (2025) and Eloundou et al. (2024). Each occupation comprises several tasks, with definitions sourced from O*NET, which offers detailed descriptions for all recognized occupations. Our first step is to determine whether a task can be replaced or enhanced by AI, a process known as task-level exposure. This is accomplished by extracting verb-noun pairs from task descriptions and the AI patent corpus. The verb-noun pairs from AI patents illustrate the functions that AI technologies can perform, whereas those from task descriptions capture the objectives and actions of human tasks.

To quantify AI exposure at the task level, we compute the cosine similarity between the $q_m \times 1$ verb–noun embedding vector of each task k in year m ($U_{k,m}$) and that of the AI patent universe in year m ($U_{\text{Patent},m}$) to quantify the degree of AI exposure at the task level. Both embedding vectors are ℓ_2-normalized. Specifically, we calculate the cosine similarity for each task k in year m:

$$\text{AI}_{k,m}^{\text{Task}} = U_{k,m}^{\top} U_{\text{Patent},m}. \tag{2}$$

After constructing the task-level AI exposure, we aggregate this information to obtain the occupation-level AI exposure ($\text{AI}_{j,m}^{\text{Occ}}$) following the same procedure as in Eisfeldt et al. (2025) and Eloundou et al. (2024). Using this occupation-level measure, we then construct the AI exposure for 3- to 6-digit NAICS industries based on the Bureau of Labor Statistics

(BLS) industry occupation survey data:

$$\mathrm{AI}^{\mathrm{Ind}}_{s,m} = \sum_{j=1}^{M_{s,m}} w_{j,s,m}\,\mathrm{AI}^{\mathrm{Occ}}_{j,m}, \qquad w_{j,s,m} = \frac{E_{j,s,m}}{\sum_{j=1}^{M_{s,m}} E_{j,s,m}}, \tag{3}$$

where $\mathrm{AI}^{\mathrm{Ind}}_{s,m}$ is the industry-level AI exposure for industry s in year m, $M_{s,m}$ is the total number of occupations in industry s and year m, and $E_{j,s,m}$ is employment of occupation j in industry s in year m. By construction, $\sum_{j=1}^{M_{s,m}} w_{j,s,m} = 1$. Since firm-level occupational data—such as those provided by Revelio Labs—are not publicly available, we aggregate our AI labor exposure measure to the 3- to 6-digit NAICS industry level, consistent with the granularity of the BLS survey data. Finally, we extrapolate industry-level AI labor exposure to the firm level using each firm's corresponding finest NAICS code (3–6 digits, depending on the industry and the survey data) and define this as our AI labor measure ($\mathrm{AI}^{\mathrm{Labor}}_{i,m}$). This aggregation should have a limited impact on our subsequent analysis, as the cross-section includes more than 400 industry groups, providing sufficient heterogeneity for a meaningful cross-sectional investigation.

Our approach to linking tasks, occupations, and firms, while conceptually similar to previous studies on labor exposure to generative AI (e.g., Eisfeldt et al., 2025 and Eloundou et al., 2024), presents several key differences. Firstly, we define AI's impact on labor more broadly, extending beyond generative AI. Instead of using ChatGPT to assess whether a task can be replaced by generative AI, we evaluate task-level exposure by examining the overlap between the AI patent universe and task descriptions. Secondly, we ensure consistency between AI product and labor exposure measures by anchoring both in the same AI patent corpus, rather than relying on a pre-defined set of keywords. Lastly, we use cosine similarity to better capture semantic relationships, as keyword frequency counts are sensitive to text length and heavily influenced by the chosen vocabulary set.

Measurement errors are inherently associated with the 10-K-based AI indicator. To mitigate these concerns, we incorporate two alternative measures of AI product exposure: the proportion of AI-related employees and the proportion of AI-related patents relative to total patents Babina et al. (2023). For AI labor exposure, we use NLP methods to define task-level AI exposure. Specifically, we employ a text-embedding approach that converts both patent texts and task descriptions into 4,096-dimensional vectors using pre-trained embedding models. The connection between tasks and AI patents is measured by the cosine similarity between their respective embedding vectors. This method captures not only verb-noun overlaps but also the contextual meaning of the text. An increasing body of literature has adopted these advanced textual analysis methods (e.g., Hampole et al., 2025 and Seegmiller et al., 2023). A detailed discussion on the implementation of text-embedding methods for estimating task-level AI exposure can be found in Section D.

3.2 Stock Price and Other Firm Characteristics

We focus on firms in the US market from 2012 to 2024. Our primary sample comprises firms at the intersection of 10-K filing data, monthly CRSP returns, and Compustat datasets. We obtain 10-K filing data utilizing the API services provided by D2V.[6] Additionally, we employ the WRDS SEC Analytics tool to map each SEC Central Index Key (CIK) to its Compustat Global Company Key (GVKEY) historically. Our analysis includes stocks listed on the NYSE, AMEX, and Nasdaq. We specifically consider ordinary common shares, identified by share codes 10 and 11, and incorporate CRSP delisting returns in our study. In line with existing literature (Shumway, 1997 and Shumway and Warther, 1999), when a stock's delisting return is unavailable, and the delisting is performance-related, we impute a return of -30% for stocks listed on the NYSE and AMEX and -55% for Nasdaq stocks. After the filtering process, we have around 41000 firm-year observations for AI exposures in our sample period.

To construct a monthly dataset for the AI exposure measure, we assume that the AI indicator for year m is updated 6 months after the fiscal year ends. For instance, if a firm's fiscal year concludes in September, the AI indicator is updated in March of the following year. Additionally, we assume the AI indicator remains constant throughout each 12-month period.

In addition to stock returns, we incorporate several firm characteristics into the regression analysis. These include market value, book-to-market ratio, operating profitability, asset growth, tangibility, stock return volatility, cash holdings (Green et al., 2017), leverage ratio, stock illiquidity, and R&D intensity. Detailed definitions of all variables can be found in Table G.8. Annual accounting data are sourced from Compustat.

Panel A of Table 1 presents the summary statistics for all variables used in the analysis. The average cosine similarity for AI product exposure is 0.16. We observe substantial heterogeneity among firms, with exposure at the 90th percentile being nearly twice that at the 25th percentile. In contrast, the AI labor exposure shows an average cosine similarity of 0.005, markedly smaller than that of product exposure, and exhibits significantly less cross-sectional variation. Panel B displays the correlations between the different AI exposure measures. The AI product and AI labor exposures are nearly uncorrelated, aligning with findings from previous research Eisfeldt and Schubert (2024), indicating that product and labor exposure to ChatGPT are largely independent.

3.3 AI Exposures: Stylized Facts

Before analyzing the relationship between AI exposure and stock prices, we discuss the characteristics represented by our AI product and labor exposure measures. Figure 2 illustrates the average AI product exposure across two-digit NAICS sectors annually. The

[6] https://sec-api.io/.

information industry (NAICS 51) shows the highest AI product exposure among all sectors. This finding aligns with expectations, as technology firms predominantly develop AI patents, and AI technologies are most commonly applied in data processing and generation. Notably, we also observe significant exposure in several non-technology industries, such as healthcare (NAICS 62) and education (NAICS 61). Conversely, industries with the lowest levels of AI product exposure include farming (NAICS 11) and mining (NAICS 21).

Next, we focus on analyzing AI labor exposure, primarily at the occupation and industry levels. Readers seeking further details on task-level analysis and the validation of our verb-noun pairing method should refer to Section C. We discuss examples of occupations with the highest AI exposure in Section C. Panel A of Figure 3 presents the average AI labor exposure across two-digit SOC codes.

The level of AI labor exposure has increased steadily over time. Two broad categories of occupations exhibit particularly high exposure: those with a high degree of task automation in lower-skilled activities (e.g., SOC 43, Administration; SOC 45, Farming; and SOC 49, Maintenance) and those heavily reliant on data processing (e.g., SOC 15, Computer; SOC 19, Social Sciences). In contrast, occupations involving extensive face-to-face interaction, such as food preparation (SOC 35) and personal care services (SOC 39), remain less exposed to AI technologies.

Panel B presents the average AI labor exposure at the industry level. Industries such as professional services and information technology (NAICS 54) display high and rising levels of AI labor exposure, reflecting their early adoption and integration of AI tools. Meanwhile, sectors that depend on personal interaction – such as retail trade (NAICS 45) and accommodation (NAICS 72) – show relatively limited AI labor exposure.

Despite the rapid advancement of AI technologies, we observe relatively limited changes in both AI labor and product exposure over time. This finding contrasts with previous studies documenting more pronounced temporal variation in AI-related vacancies (e.g., Acemoglu et al., 2022b), which suggests AI exposure should vary over time. One possible explanation is that most firms maintain consistent business descriptions and task definitions, resulting in limited time-series variation when exposure is aggregated at the two-digit industry level. This limitation, however, does not affect our analysis, which focuses primarily on cross-sectional differences. Section B provides examples of firms with high and evolving AI product exposure in non-technology industries, where we observe significant time-series variation for some companies. Moreover, Table 1 confirms substantial firm-level heterogeneity in both AI product and labor exposure.

4 Baseline Results

4.1 Portfolio Sorts

We begin by conducting a portfolio-sorting analysis. At the end of each month t, we sort all firms into decile portfolios based on their AI exposures for month t. We then maintain both equal-weighted and value-weighted portfolios throughout month $t+1$. We calculate the average return, as well as three- to six-factor alphas, and assess the return performance of long-short portfolios formed by taking positions in the highest and lowest deciles. The results are presented in Table 2.

Panel A reports portfolio returns based on AI product exposure, revealing positive returns and alphas in value-weighted portfolios but not in equal-weighted portfolios. This suggests that the AI product premium is likely concentrated among large firms. This observation aligns with findings by Acemoglu et al. (2023), who noted that advanced technology adoption, including AI, is primarily concentrated within large companies.

Panel B presents results for AI labor exposure. Although the long-short portfolios yield positive returns, these results are not statistically significant. Unlike the product exposure results, equal-weighted portfolios demonstrate higher expected returns than value-weighted portfolios. This implies that the AI labor premium may be more pronounced among smaller firms.

Figure 4 plots the cumulative returns of long-short portfolios sorted by AI labor and product exposures. A distinct time-varying pattern emerges in the AI product premium. Before 2016, the cumulative return of the AI product-based long-short portfolio hovered near zero. It turns positive post-2016, declines with the onset of 2021, and then rises sharply after 2023. The positive returns are more pronounced in value-weighted portfolios compared to equal-weighted ones (Table 2). Notably, significant positive movements after 2023 occur exclusively in the value-weighted portfolios.

In examining AI labor exposure, we also observe a pattern that varies over time. For value-weighted portfolios, returns remain near zero. In contrast, equal-weighted portfolios show modestly positive returns before 2020, followed by a sharp decline during the COVID-19 period. The portfolio-sorting results overall indicate a strong positive premium associated with AI product exposure and a weaker positive premium linked to AI labor exposure. The time-varying nature of these premiums suggests multiple underlying factors are influencing the observed return dynamics.

4.2 Regression Results

Next, we examine the AI premium using the cross-sectional regression of Fama and MacBeth (1973), estimated for each month t:

$$R_{i,t} = \alpha_t + \beta_t^{AIMeasure}\text{AI Measure}_{i,t-1} + \delta_t\text{Controls}_{i,t-1} + \epsilon_{i,t}, \tag{4}$$

where $R_{i,t}$ is the return of company i for month t, AI Measure$_{i,t-1}$ is the AI exposure information available to the investor at month $t-1$. The parameter $\beta_t^{AIMeasure}$ measures the expected return differences in each month t in response to changes in AI exposure information. If AI Measure$_{i,t}$ is standardized cross-sectionally, then $\beta_t^{AIMeasure}$ can also be interpreted as the return on a long–short portfolio that takes a long position in high-AI firms and a short position in low-AI firms (Fama and French, 2020). The sample period spans from January 2012 to December 2024. Controls$_{i,t-1}$ are a set of controls defined in Table G.8. All independent variables are cross-sectionally winsorized at the [1%, 99%] level and standardized to have zero mean and standard deviation of unity. We then test whether the time-series averages of the regression coefficients in equation (4) are statistically significant. The standard errors are based on the Newey and West (1987) adjustment with a lag of 12 months. Companies with a closing price of less than 5 USD in month $t-1$ are excluded from the cross-sectional regression.

Table 3 reports the regression results. Columns (1) to (3) present estimates for the full sample period. The coefficient on AI product exposure is positive and statistically significant (0.07, t-statistics = 2.00). In contrast, the premium associated with AI labor exposure is insignificant. These findings align with previous evidence concerning the market valuation effects of AI investment. For example, Babina et al. (2024a) document a positive relationship between firm market value and new AI investments during 2010–2018, while Mishra et al. (2022) find that firm-level AI attention is positively related to firm value.

In Columns (4) to (6), we limit the sample period to conclude in December 2019, representing the pre-COVID era, and re-estimate the model. Within this subsample, we observe a stronger relationship between AI product exposure and stock returns, evidenced by a coefficient of 0.08 (t-statistics= 2.03). Although the AI labor premium increases, it remains statistically insignificant. Subsequently, we conduct cross-sectional regressions for each month t and illustrate the cumulative coefficients in Figure 5. The AI premium demonstrates significant time variation; recent negative shocks may have diminished its average effect over the entire sample period. In untabulated results, when we extend the sample period from 2012 to the end of 2020, we find an even stronger AI product premium, with a coefficient of 0.11 (t-statistic = 2.55) at its peak.

In both figures, we observe strong positive shocks in AI product premiums beginning in 2018. Table G.15 presents a cross-sectional regression conducted from January 2012 to December 2017. During this sub-period, the coefficient for product exposure becomes insignificant, with an average premium of 0.04 (t-statistics= 1.03) , which is lower than the premium for the entire 2012-2019 period. This suggests that AI product premiums were insignificant in the early years, indicating a potential slow realization process for the premium.

Our findings suggest that the AI premium is multifaceted. Primarily, its realization has occurred through the AI product channel over the past decade, albeit gradually and with delay. Historically, the product channel has dominated; however, the AI labor channel started showing effects after mid-2022. This shift reflects growing market interest in the potential impact of AI technologies on firm labor. Notably, a sharp decline was observed around 2020. We explore these dynamics in greater detail in the next section.

4.3 Beyond Tech Companies

As depicted in Figure 2, the sectors with the highest levels of both product and labor exposure are predominantly technology-related. This is logical since technology firms are known for actively embracing and applying AI. Furthermore, the "Magnificent Seven" stocks – the seven largest companies leading recent market index performance – are at the forefront of AI technology development.[7] Therefore, a plausible explanation for the positive premium is the overall boom in the technology sector.

To test this conjecture, we re-estimated the regressions without including technology firms. We excluded companies with NAICS two-digit codes of 51 and 54, as well as the seven Magnificent Seven stocks. The regression results are reported in Table G.14. Notably, the coefficient on AI product exposure for the full-sample period increased when technology firms were excluded (0.1, t-statistic = 2.31). This finding implies that although AI innovations primarily originate from the technology sector, the observed AI premium is not exclusively driven by this sector. Regarding the AI labor premium, we found it to be negative for the full-sample period, indicating that technology-related companies might have been more resilient during the COVID-19 period.

Next, we present the cumulative value of the AI product and labor premium in Figure G.10. In each graph, we plot both the AI premium with the full-sample and excluding technology-related companies. First, consistent with the full-sample regression results, the AI product premium is even higher when we exclude technology-related companies. In addition, both premiums display similar patterns, even though approximately 12% of total observations are removed. Second, the decline during the COVID-19 period is much smaller for the full-sample, confirming the conjecture that technology-related stocks performed better during the pandemic.

[7] These companies include: Apple (AAPL), Amazon (AMZN), Alphabet (GOOG), Meta Platforms (META), Microsoft (MSFT), Nvidia (NVDA), and Tesla (TSLA). See, for example, discussions at https://www.investopedia.com/magnificent-seven-stocks-8402262.

5 Chasing the AI Premium

The findings from the previous section partially support recent evidence indicating that AI investments are acknowledged and valued by capital markets. In this section, we further examine the underlying mechanism responsible for the observed premium.

5.1 Slow Diffusion of AI Technology

A common view in the literature is that the diffusion of new technology is slow (e.g.,Kalyani et al., 2025 , Manuelli and Seshadri, 2014 and Atkeson and Kehoe, 2007). In our case, the AI product premium during early periods is not significant, which indicates that the same slow diffusion might apply to AI technology. To examine the conjecture, we study the predictability at a longer investment horizon. We perform the cross-section regression with 1, 3, 6, 12, 18, 24 and 36-month returns.

Figure 6 illustrates the coefficient and its 95% confidence interval. Panel A presents the analysis of AI product exposure. In the pre-COVID period (left graph), predictability is significant across all return horizons, increasing with the length of the horizon. For the full-sample period (right graph), predictability at the mid-term horizon decreases compared to the pre-COVID period. This aligns with observations in Figure 5, where negative shocks were evident after 2021. Nevertheless, predictability at longer horizons remains positive and significant. Regarding AI labor exposure, results show positive but insignificant impacts at all horizons, suggesting that the influence of AI labor during the sample period is limited.

The results discussed have several implications. If the impact of AI technology were solely a short-term phenomenon, long-term predictability would approach zero. However, our findings indicate both long-term and short-term predictability, as well as mid-term reversals. This suggests the presence of fundamental impacts – such as those on company performance as documented in the literature – and behavioral factors. Additionally, in the full-sample analysis from 2012 to 2024, we observe a sharp increase in predictability beyond 24 months. This indicates that significant market realization of product impact may take, on average, at least 2 years. These results align with findings in Babina et al. (2024a), which indicate it takes at least 2 to 3 years for firms to realize the benefits from AI investments.

5.2 Is AI Premium a Mispricing?

As illustrated in Figure 5, there are negative shocks following 2021. If investors value AI purely for its capacity to enhance firm productivity or innovation – a "risk-based" interpretation – there would be no reversals. Therefore, does the decline in the AI product premium after 2021 reflect shifts in market attention or a reassessment of the expected benefits of AI technologies? A plausible hypothesis is that part of the negative premium stems from investors adjusting their expectations. Some of the observed AI premiums may, in fact, be due to mispricing rather than fundamental risk compensation. Leung et al. (2020) find that the pricing of R&D in the cross-setion of returns represents a mispricing narrative.

To test this conjecture, we conduct two complementary analyses. First, we investigate whether fluctuations in the AI premium correlate with changes in market attention related to AI. We download the monthly Google Trends Index using the keywords "Artificial Intelligence" and "AI". We estimate an AR(1) model for the monthly changes in the index, utilizing the residuals to capture unexpected shocks in AI attention. We then interact this attention measure with firm-level AI exposure to predict returns. We anticipate that extreme shocks to the AI premium will coincide with heightened AI attention – especially during recent periods of the AI boom.

Table 4 presents the results of pooled panel regressions that include interaction terms for AI product exposure and AI labor exposure. The interaction term for AI product exposure is positive and statistically significant, indicating that periods of heightened AI attention correspond with a stronger AI product premium. Conversely, the interaction between AI labor exposure and attention is negative and significant. This suggests that investors are inclined to sell firms with high AI labor exposure during times of increased AI attention. This tendency may explain the negative premium observed during the COVID-19 period.

We investigate whether institutional investors adjust their holdings in AI-related companies following earnings announcements. To explore this, we obtain quarterly SEC 13-F filing data and compute the percentage of outstanding shares held by institutional investors for each firm.[8] We then assess the relationship between firms' exposure to AI products and both the level of institutional holdings and changes in these holdings each quarter. This analysis is conducted using the following pooled OLS regressions:

$$IO_{i,q+1} = \alpha + \beta_1 AI_{i,q}^{Labor} + \beta_2 AI_{i,q}^{Product} + \gamma Controls_{i,q} + \epsilon_{i,q}. \tag{5}$$

The control variables are the same as in (4). We consider both the level of institutional holding and quarterly changes in institutional holdings. We expect that firms with higher AI product exposure exhibit higher institutional ownership. Furthermore, if there are behavioral factors, we should observe negative adjustments in holdings following positive shocks, which may help explain the decline in the AI product premium observed after 2021.

Figure 7 plots the estimated coefficients and their 95% confidence intervals over time.

[8]The 13-F filing data are retrieved via the SEC API: https://sec-api.io/docs/query-api/13f-institutional-ownership-api. Appendix E provides a detailed description of how firm-level institutional ownership data are extracted from the filings.

Panel A illustrates the relationship between AI product exposure and institutional holdings, while Panel B displays the relationship between AI labor exposure and institutional holdings. First, we find a positive relationship between institutional ownership and AI exposure, which has strengthened in recent periods. Second, consistent with our expectations, we observe a negative association between changes in institutional holdings and the degree of AI exposure – in both labor and product measures – during certain periods. These results suggest that, while AI technologies impact firm fundamentals and valuations, behavioral factors significantly shape the observed AI premiums.

5.3 Task Importance and Firm Size

In the previous section, we demonstrated that the negative AI labor premium might reflect investors' revised expectations. As illustrated in Figure 5 and Table 3, we observed only a weak positive relationship between AI labor exposure and asset returns. This suggests that the labor impact of AI is not fully incorporated into capital market valuations. This finding partially contrasts with prior studies on AI's labor effects, which often highlight positive outcomes for firm performance. For example, Chen and Wang (2024) show that AI innovations, measured by the number of AI-related patents, result in higher firm valuations, especially for companies with lower external hiring costs.

The key question is what our measure of AI labor exposure precisely captures. A higher level of AI labor exposure indicates that a company's workforce is more *prone* to being replaced or supported by AI technology. Ideally, companies with occupations vulnerable to AI technology (indicated by high AI labor exposure) actively seek and implement AI solutions. This would lead to significant effects on firm productivity, which investors would recognize and subsequently reflect in asset prices. However, this scenario is less common in the early stages, partly explaining the weak labor premium.

The actual effect of AI on labor within a company depends on three factors: 1) how the company applies AI technology to improve labor efficiency; 2) how other firms use AI technology for labor efficiency improvement; and 3) how easily AI technology can be applied (e.g., ChatGPT versus other commercial AI applications). The situation becomes complex when all three factors are considered, and these factors interact with each other to produce asset price impacts, making it empirically challenging to isolate each factor.

We can still derive insights into the mechanism through which AI influences labor efficiency: either through the substitution or augmentation of labor. One explanation for our findings is that AI technologies exert both substitution and augmentation effects, which may counterbalance each other across firms, or there might be variation in these effects. For example, Eisfeldt et al. (2025) found that although generative AI generally enhances firm value, the substitution effect outweighs the augmentation effect. Conversely, Hampole et al. (2025) observed muted effects of AI on employment due to offsetting forces: occupations highly exposed to AI face relatively lower demand compared to less exposed occupations. However, the increase in firm productivity ultimately elevates overall employment across all occupations.

Following Eisfeldt et al. (2025), we account for the differing importance of tasks when constructing our AI labor exposure measure. In the O*NET framework, tasks are classified as either core or supplemental to an occupation, with core tasks holding greater significance. In our main analysis, we weight core tasks at 0.7 and supplemental tasks at 0.3 when aggregating task-level AI exposure to the occupational level. Additionally, we develop two alternative measures with different weighting schemes to account for task importance: the AI_{Labor}^{Supp}, which assigns a weight of 1 to supplemental tasks, and AI_{Labor}^{Core}, which assigns a weight of 1 to core tasks.

Figure 8 plots the cumulative returns associated with the three AI labor premium measures. All three exhibit broadly similar patterns over time; however, the premium based on supplemental tasks is higher than that of core tasks. This suggests that, in the earlier years, the benefits of AI were primarily realized through labor augmentation, where AI technologies played a supportive role. Beginning in 2023, we observe a sharp increase in the AI_{Labor}^{Core} premium, indicating a shift from supportive to more central applications of AI in core occupational tasks – a pattern consistent with the findings of Eisfeldt et al. (2025).

Table 5 reports the average returns for the pre-COVID period, alongside the corresponding alphas from spanning regressions of long-short portfolios sorted by the different AI labor exposure measures. Consistent with the patterns observed in Figure 8, we find that the equal-weighted portfolios based on supplemental tasks yield weak positive alphas, whereas the average return associated with core tasks is nearly zero. Notably, the average return for the equal-weighted portfolio sorted on AI_{Supp}^{Core} is five times higher than that of AI_{Labor}^{Core}.

We further investigate the potential reasons for the negative aggregate premium observed during the COVID-19 period. One possible explanation involves the disproportionate impact of the pandemic on small businesses (Bartik et al., 2020). To examine this hypothesis, we focus on data from January to December 2020 and conduct a predictive panel regression of returns based on AI labor exposure and its interaction with firm size. The findings, detailed in Table 6, reveal that the interaction term is both positive and significant. This indicates that the negative association between returns and AI labor exposure is most pronounced for smaller firms and diminishes as firm size increases. Additionally, the AI labor premium is negative across all three types of AI labor exposures. Therefore, the negative premium is likely attributable to the poor performance of small firms with substantial AI labor exposures.

5.4 The "AI Risk"

While most previous studies emphasize the benefits of AI technology, the potential risks and uncertainties associated with its advancement are often overlooked. For instance, firms with substantial AI exposure are more vulnerable to AI-related regulatory risks. A notable example is the temporary ban of ChatGPT in early 2023. Additionally, high-AI-dependent firms may encounter elevated cybersecurity risks due to their reliance on data-driven operations. Conversely, firms with low AI exposure risk technological obsolescence, a condition referred to as the "risk of laggards". By failing to adopt AI innovations, these firms may lose product competitiveness.

To examine "AI risk", we extract the Risk Factors section (Item 1A) from firms' 10-K reports and apply the procedure used to construct the AI product exposure measure. Specifically, we calculate the overlap between the AI patent universe and each firm's risk factor disclosures to obtain an AI risk exposure measure. As reported in Table 1, the correlation between AI risk exposure and AI product exposure is relatively low. This suggests that the AI risk measure captures distinct information not contained in product exposure

What exactly is captured by the AI-related risk exposure extracted from the Risk Factors section? To provide a comprehensive overview of the key themes associated with AI risk, we employ the Latent Dirichlet Allocation (LDA) model proposed by Blei et al. (2003). This model identifies latent topics by estimating discrete probability distributions over words for each topic and inferring per-document topic distributions. We adopt an enhanced version known as Local-LDA, following Brody and Elhadad (2010), in which each sentence, rather than the entire document, is treated as a separate document. Lu et al. (2011) compare different topic-modeling settings and find that Local-LDA performs best in capturing distinct and meaningful topics in consumer reviews.

We construct two sets of document groups to highlight nuances in AI risk exposure. First, each fiscal year, we rank firms based on their AI risk exposure, classifying them as either "high-AI-risk" or "low-AI-risk" using the 90th and 10th percentiles as cutoffs. Second, we independently rank firms by AI product exposure in the same fiscal year, again classifying them as "high-AI-product" or "low-AI-product" using the same cutoffs. We focus on two groups of particular interest: firms that are both high-AI-product and high-AI-risk (Group 1) and firms that are low-AI-product but high-AI-risk (Group 2). For each fiscal year, we then apply the Local-LDA topic model. Following prior studies such as Bao and Datta (2014) and Lopez Lira (2019), we set the number of topics to be 25.

The LDA analysis produces topics represented as a bag-of-words and their relative importance, with importance scores summing to one across all topics. To generate human-interpretable topic labels, we use a state-of-the-art large language model, DeepSeek-v3.1-

pro, to classify and interpret each bag-of-words. This labeling process adheres to the topic taxonomy developed by Bao and Datta (2014), with supplementary categories addressing product-specific risks. Detailed definitions of risk are provided in F.

Tables G.16 (Group 1) and G.17 (Group 2) present the top ten topics over time from 2016 to 2023. Overall, the AI risk captured for high-AI-product firms (Group 1) is primarily associated with new product risks, cybersecurity threats, intellectual property disputes, challenges related to rapid innovation, and the high costs of R&D. In contrast, for low-AI-product firms (Group 2), both cybersecurity risk and product risk are of lower importance than for Group 1, which indicates that adopting the new technology does involve risks associated with AI. Table G.18 provides representative excerpts from firms' risk disclosures. The first three rows correspond to Group 2 firms, all of which highlight concerns about cybersecurity and the risk of lagging behind in AI adoption. The last three rows feature examples from Group 1 firms, notably Recursion Pharmaceuticals, which emphasizes legal and product-related risks in AI-driven drug discovery. Overall, the AI risk factor comprises two distinct dimensions of technological threat: risks from adopting and using AI, and risks from not adopting it.

We conduct two empirical tests using the constructed AI risk measure. First, we examine whether AI risk is reflected in the cross-section of stock returns. Second, we investigate whether AI product exposure interacts with AI risk exposure in explaining asset returns. Our hypothesis posits that higher AI risk should correlate with a positive risk premium. Additionally, we propose that the interaction between AI risk and AI product exposure should amplify this effect, as investors require additional compensation for enduring greater technological and regulatory uncertainty.

Table 7 presents the regression results for three subperiods: 2012–2024 (full), 2012–2019 (pre-COVID), and 2023–2024 (recent). The analysis includes both the level of AI risk exposure and its interaction with the AI product measure. Our findings reveal a significant positive premium for AI risk during the full-sample period (0.076, t-statistics = 2.39). This premium is even higher in the pre-COVID period (0.085, t-statistics = 2.41). Figure 9 illustrates the cumulative value of the AI risk premium. The pattern closely aligns with that of the AI product premium, showing positive shocks from 2018 to late 2023 and a renewed increase starting in 2024

The interaction term between AI risk and AI product exposure, although positive, is insignificant for the full-sample period and is nearly zero for the pre-COVID period (0.009, t-statistics = 0.26). This indicates that while AI-related risk seems to be priced in the market, it does not have a direct relationship with the AI product premium. This finding makes intuitive sense: as the development and commercial adoption of AI technologies were still in their early stages during much of the sample period, increased AI product exposure did not necessarily lead to heightened AI-related risk. Another possible explanation is that the risk captured by our AI risk measure – particularly during the pre-COVID period – may not pertain to AI but rather to general data-related risks, such as cybersecurity risk, which could explain why the interaction term is almost zero during the pre-COVID period.

That being said, the interaction term for the period from January 2023 to December 2024 is 0.13 ($t = 1.76$), which is significantly higher than the results for earlier periods. This finding indicates that in recent years, AI risk has increasingly influenced the explanation of the AI product premium. Overall, while AI-related risks are priced into capital markets, we find limited evidence that these risks serve as the primary driver of the observed results over the entire sample period.

5.5 AI, Digitalization and Firm-Level Risks

The application of AI technologies demands substantial digital infrastructure, including data centers and high-capacity computing systems. Prior studies have demonstrated the positive impact of digital transformation on firm value. For instance, Tambe (2014) shows that adopting big data technologies enhances firm productivity, while Chen et al. (2019) find that FinTech innovations are positively valued by investors. Therefore, the AI premium identified in our analysis may partly reflect the benefits of digitization – namely, the value gains arising from transforming firm operations through digital technologies.

To test this conjecture, we construct a digitalization exposure measure based on the business descriptions in firms' 10-K reports. Following Chen and Srinivasan (2023), we compile a keyword list related to big data, cloud computing, and digitization to capture other forms of digital technologies.[9] We then estimate the cosine similarity between this keyword list and each firm's business description to quantify the degree of digitalization. Columns (1) and (2) of Table G.19 present the cross-sectional regression results. The AI product premium remains positive and significant after controlling for digitalization exposure. Although the interaction term between AI product exposure and digitalization is positive, it is not statistically significant, suggesting that the AI product premium is not primarily driven by the general benefits of digitalization.

Next, we investigate whether the observed AI premium is driven by other forms of firm-level risk. We incorporate three risk measures from Hassan et al. (2019)—non-political risk, political risk, and overall risk—into the Fama-MacBeth regression framework.[10] Columns (3) – (8) of Table G.19 presents the results. The magnitude and significance of the AI premium remain largely unchanged after controlling for these additional risk factors. This finding suggests that the AI premium is not influenced by broader firm-level risks.

[9] big data, data science /mining, data lake, devops, digital/twin, edge computing, cloud platforms/enablement, virtual machines, digitization, digital transformation/revolution/strategy/marketing, business/customer/operating intelligence

[10] Data are obtained from: https://www.firmlevelrisk.com/home.

6 Robustness and Further Results

In this section, we explore several alternative settings in addition to the current discussion. Additional empirical results are provided in the Appendix G.

6.1 Different AI Technologies

If AI technologies confer a positive product premium due to their potential benefits to a company, it is reasonable to conjecture that these benefits vary according to the type of technology. For example, Chen and Wang (2024) classifies AI technologies based on their cognitive capabilities into seven types: inference, decision-making, learning, creativity, language, perceptual-motor, and engagement. They find that different types of AI innovations have different impacts on the company. We follow Giczy et al. (2021) to further classify the AI patent universe into eight categories: machine learning, knowledge processing, speech, AI hardware, evolutionary computation, natural language processing, computer vision, and planning/control.

We classify technologies to investigate whether different AI technology applications have varying impacts on products. To achieve this, we integrate AI patent classification data with AI product exposure and adjust the portfolio-sorting process accordingly. Initially, we estimate the technology-specific exposure from AI patents.

$$\text{Tech Exposure}_{i,k} = \frac{\text{Number of Patents Related to Technology } k_i}{\text{Total Number of Patents}_i}, \tag{6}$$

Next, the AI product exposure related to technology k is

$$\text{AI Product Exposure}_{i,k,m} = \text{Tech Exposure}_{i,k} \times \text{AI}_{i,m}^{\text{10K Filling}}. \tag{7}$$

Table G.20 presents data on the number of companies with non-zero Tech Exposure from 2011 to 2023, along with the average value of Tech Exposure for each year. Among the technology categories, "planning and control" boasts the highest percentage and number of companies, while "speech"registers the lowest. Over the years, "machine learning" exhibits the highest growth rate, whereas "speech" remains the least developed category.

Table G.21 presents the portfolio sort results from 2012 to 2024. There is heterogeneity across technology categories: computer vision, speech, and NLP yield positive returns, while machine learning and planning/control exhibit insignificant returns. Notably, the average return from computer vision is almost double that of machine learning. Consistent with the results in Table 2, significant returns are observed only in value-weighted portfolios. This suggests that the realization of the AI product premium primarily occurs in large companies, as these firms typically have more resources and incentives to implement AI innovations.

6.2 Alternative AI Exposure Measures

The primary challenge in quantifying AI adoption is measurement error. All measures discussed in the literature serve as indirect approximations of AI adoption, which remains an unobserved intangible asset. Firms might omit reporting their use of AI technology in 10-K filings if the technology is solely employed for internal management or is not a central business focus. Consequently, text mining-based measures also exhibit considerable noise. A higher level of cosine similarity does not necessarily reflect a greater emphasis on AI; additionally, the accuracy of this measure depends on the scope of the patent universe.

To assess the robustness of our results concerning the 10-K filings and the potential influence of measurement error, we employ two supplementary indicators for AI product exposure. Firstly, we consider the percentage of AI employees from Babina et al. (2024a). The AI exposure of company i in year m is approximated by the following indicator:

$$\mathrm{AI}_{i,m}^{\mathrm{Employee}} = \frac{\text{Number of AI-related Employees}}{\text{Total Number of Employees}}, \tag{8}$$

where an employee is classified as AI-related based on skill-level AI-relatedness in her resume. A higher percentage of AI-related employees indicates a higher level of AI adoption within the company.[11]

The second variable for the AI product exposure is based on the percentage of AI patents and grants:

$$\mathrm{AI}_{i,m}^{\mathrm{Patent}} = \frac{\text{Total Number of AI-related Patents}}{\text{Total Number of Patents}}. \tag{9}$$

We use the patent-firm mapping data from Kogan and Papanikolaou (2014).[12] If a company incorporates AI technology into its production, there will be new AI-related patents, and $\mathrm{AI}_{i,m}^{\mathrm{Product}}$ will be higher.

For the AI labor measure, instead of using the verb-noun matching strategy, we use the text-embedding technique to convert patent text and task description text into 4096-dimensional vectors and estimate the cosine similarity between task and patent embeddings to measure the task-level AI exposure. Table 1 shows the summary statistics and correlations between alternative AI measures.

We first present in Figure G.11 the cumulative value of the AI premium from different AI exposures. All three measures exhibit similar patterns: an increase before 2021, a decrease after the beginning of 2021, and a regain of positive shocks starting from 2023. Turing to the AI labor measure (lower panel), while AI_{labor}^{NLP} has a much larger size than AI_{labor}, the two measures yield quite similar patterns, indicating that the verb-noun

[11]We refer interested readers to Section 4 of the original paper for a detailed description of how the firm-level indicator is constructed.

[12]Updated database: https://github.com/KPSS2017/Technological-Innovation-Resource-Allocation-and-Growth-Extended-Data

strategy works well in capturing the task level AI exposure.

Next, we present cross-sectional regression results in Table G.22, covering data up to December 2020. All measures of AI products generate positive premiums. Therefore, although different measures may include some level of measurement error, our findings remain consistent.

7 Conclusion

The rapid development of AI technology in recent years has captured the attention of the investment community. This paper offers a detailed investigation into the diffusion and pricing of AI technology within the US stock market from 2012 to 2024. By creating novel, unified metrics of AI exposure that distinctly assess its effects on products and labor, we document a complex and dynamic relationship between AI technology and the capital market.

We identified a significant positive premium associated with AI product exposure in firms that integrate AI into their core products and services. However, this premium became observable only after 2018. In contrast, the premium for AI labor exposure – referring to the automation and augmentation of internal tasks – was generally weak and insignificant throughout most of our sample period. Moreover, the pricing of AI technology in the capital market is a slow and gradual process influenced by multiple forces. The anticipated benefits of AI adoption were not immediately reflected in stock prices during the early years.

We explore the economic mechanisms underlying these patterns. As technology becomes more widespread, the positive product premium is slowly integrated into prices, coinciding with the clarification of its commercial applications. However, we also identify a behavioral mispricing component. Notably, the AI premium related to product exposure strongly correlates with market-wide AI attention, such as Google search trends. This indicates that investor attention and speculative trading significantly influence the market, particularly during periods of technological hype. The attention-driven component contributes to understanding the subsequent negative corrections seen during the COVID period.

Furthermore, we demonstrate that AI technologies designed to automate supplemental tasks are viewed more positively than those intended to replace core occupational functions. Additionally, we discover that AI-related risks, such as regulatory and cybersecurity concerns identified in corporate filings, are independently priced in the cross-section, though they do not influence the core product premium. Our results remain robust across various tests, including the use of alternative measures of AI exposure, controls for other digital technologies, and the exclusion of the technology-related sector.

In conclusion, although the long-term effects suggest a positive valuation for firms that successfully implement AI, particularly through product innovation, the path to achieving this positive outcome involves multiple factors. Our study highlights that technological adoption is not a singular event but rather a gradual process. The stock market's valuation of such adoption reflects the complex interplay between rational expectations and shifting investor narratives.

References

Acemoglu, D., May 2024. The Simple Macroeconomics of AI.

Acemoglu, D., Anderson, G., Beede, D., Buffington, C., Childress, E., Dinlersoz, E., Foster, L., Goldschlag, N., Haltiwanger, J. C., Kroff, Z., Restrepo, P., Zolas, N. J., Nov. 2022a. Automation and the Workforce: A Firm-Level View from the 2019 Annual Business Survey. Working Paper w30659, National Bureau of Economic Research.

Acemoglu, D., Anderson, G., Beede, D., Buffington, C., Childress, E., Dinlersoz, E., Foster, L., Goldschlag, N., Haltiwanger, J., Kroff, Z., Restrepo, P., Zolas, N., May 2023. Advanced Technology Adoption: Selection or Causal Effects? AEA Papers and Proceedings 113, 210–214.

Acemoglu, D., Autor, D., Hazell, J., Restrepo, P., Apr. 2022b. Artificial Intelligence and Jobs: Evidence from Online Vacancies. Journal of Labor Economics.

Acemoglu, D., Restrepo, P., Jun. 2020. Robots and Jobs: Evidence from US Labor Markets. Journal of Political Economy 128 (6), 2188–2244.

Agrawal, A., Gans, J. S., Goldfarb, A., 2024. Artificial intelligence adoption and system-wide change. Journal of Economics & Management Strategy 33 (2), 327–337.

Atkeson, A., Kehoe, P. J., Mar. 2007. Modeling the Transition to a New Economy: Lessons from Two Technological Revolutions. American Economic Review 97 (1), 64–88.

Babina, T., Fedyk, A., He, A., Hodson, J., 2024a. Artificial intelligence, firm growth, and product innovation. Journal of Financial Economics 151, 103745.

Babina, T., Fedyk, A., He, A. X., Hodson, J., Jun. 2023. Firm Investments in Artificial Intelligence Technologies and Changes in Workforce Composition.

Babina, T., Fedyk, A., He, A. X., Hodson, J., Nov. 2024b. Artificial Intelligence and Firms' Systematic Risk.

Bao, Y., Datta, A., Apr. 2014. Simultaneously Discovering and Quantifying Risk Types from Textual Risk Disclosures. Management Science.

Bartik, A. W., Bertrand, M., Cullen, Z., Glaeser, E. L., Luca, M., Stanton, C., Jul. 2020. The impact of COVID-19 on small business outcomes and expectations. Proceedings of the National Academy of Sciences 117 (30), 17656–17666.

Bertomeu, J., Lin, Y., Liu, Y., Ni, Z., Feb. 2025. The impact of generative AI on information processing: Evidence from the ban of ChatGPT in Italy. Journal of Accounting and Economics, 101782.

Blei, D. M., Ng, A. Y., Jordan, M. I., Mar. 2003. Latent dirichlet allocation. J. Mach. Learn. Res. 3 (null), 993–1022.

Brody, S., Elhadad, N., Jun. 2010. An unsupervised aspect-sentiment model for online reviews. In: Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics. HLT '10. Association for Computational Linguistics, USA, pp. 804–812.

Brynjolfsson, E., Mitchell, T., Rock, D., May 2018. What Can Machines Learn, and What Does It Mean for Occupations and the Economy? AEA Papers and Proceedings 108, 43–47.

Brynjolfsson, E., Rock, D., Syverson, C., Jan. 2021. The Productivity J-Curve: How Intangibles Complement General Purpose Technologies. American Economic Journal: Macroeconomics 13 (1), 333–372.

Chen, M. A., Wang, J. X., Apr. 2024. Displacement or Augmentation? The Effects of AI Innovation on Workforce Dynamics and Firm Value.

Chen, M. A., Wu, Q., Yang, B., Apr. 2019. How Valuable Is FinTech Innovation? The Review of Financial Studies 32 (5), 2062–2106.

Chen, W., Srinivasan, S., Mar. 2023. Going digital: Implications for firm value and performance. Review of Accounting Studies.

Cockburn, I. M., Henderson, R., Stern, S., Mar. 2018. The Impact of Artificial Intelligence on Innovation. Working Paper 24449, National Bureau of Economic Research.

Eisfeldt, A. L., Schubert, G., Oct. 2024. AI and Finance.

Eisfeldt, A. L., Schubert, G., Zhang, M. B., Taska, B., May 2025. Generative AI and Firm Values.

El Moujahid, O., Murtinu, S., Sekerci, N., Aug. 2023. Artificial Intelligence and Firm Performance during COVID-19. Working Paper.

Eloundou, T., Manning, S., Mishkin, P., Rock, D., Jun. 2024. GPTs are GPTs: Labor market impact potential of LLMs. Science.

Fama, E. F., French, K. R., 2020. Comparing cross-section and time-series factor models. The Review of Fianncial Studies 33 (5), 1891–1926.

Fama, E. F., MacBeth, J. D., 1973. Risk, Return, and Equilibrium: Empirical Tests. Journal of Political Economy 81 (3), 607–636.

Felten, E., Raj, M., Seamans, R. C., Aug. 2019. The Effect of Artificial Intelligence on Human Labor: An Ability-Based Approach. Academy of Management Proceedings.

Felten, E. W., Raj, M., Seamans, R., May 2018. A Method to Link Advances in Artificial Intelligence to Occupational Abilities. AEA Papers and Proceedings 108, 54–57.

Fotheringham, D., Wiles, M. A., Jul. 2023. The effect of implementing chatbot customer service on stock returns: An event study analysis. Journal of the Academy of Marketing Science 51 (4), 802–822.

Gârleanu, N., Panageas, S., Yu, J., 2012. Technological Growth and Asset Pricing. The Journal of Finance 67 (4), 1265–1292.

Giczy, A. V., Pairolero, N. A., Toole, A., Aug. 2021. Identifying Artificial Intelligence (AI) Invention: A Novel AI Patent Dataset.

Green, J., Hand, J. R. M., Zhang, X. F., Dec. 2017. The Characteristics that Provide Independent Information about Average U.S. Monthly Stock Returns. The Review of Financial Studies 30 (12), 4389–4436.

Greenwood, J., Jovanovic, B., 1999. The Information-Technology Revolution and the Stock Market. The American Economic Review 89 (2), 116–122.

Hampole, M., Papanikolaou, D., Schmidt, L. D., Seegmiller, B., Feb. 2025. Artificial Intelligence and the Labor Market.

Hassan, T. A., Hollander, S., van Lent, L., Tahoun, A., Aug. 2019. Firm-Level Political Risk: Measurement and Effects. The Quarterly Journal of Economics 134 (4), 2135–2202.

Hirshleifer, D., Hsu, P.-H., Li, D., Mar. 2013. Innovative efficiency and stock returns. Journal of Financial Economics 107 (3), 632–654.

Hoberg, G., Phillips, G., Oct. 2016. Text-Based Network Industries and Endogenous Product Differentiation. Journal of Political Economy 124 (5), 1423–1465.

Honnibal, M., Johnson, M., Sep. 2015. An Improved Non-monotonic Transition System for Dependency Parsing. In: Màrquez, L., Callison-Burch, C., Su, J. (Eds.), Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, Lisbon, Portugal, pp. 1373–1378.

Hsu, P.-H., 2009. Technological innovations and aggregate risk premiums. Journal of Financial Economics 94 (2), 264–279.

Kalyani, A., Bloom, N., Carvalho, M., Hassan, T., Lerner, J., Tahoun, A., Apr. 2025. The Diffusion of New Technologies. The Quarterly Journal of Economics 140 (2), 1299–1365.

Kelly, B., Papanikolaou, D., Seru, A., Taddy, M., Sep. 2021. Measuring Technological Innovation over the Long Run. American Economic Review: Insights 3 (3), 303–320.

Kogan, L., Papanikolaou, D., 2014. Growth Opportunities, Technology Shocks, and Asset Prices. The Journal of Finance 69 (2), 675–718.

Korinek, A., Suh, D., Mar. 2024. Scenarios for the Transition to AGI.

Laitner, J., Stolyarov, D., Sep. 2003. Technological Change and the Stock Market. American Economic Review 93 (4), 1240–1267.

Lee, C. M. C., Sun, S. T., Wang, R., Zhang, R., 2019. Technological links and predictable returns. Journal of Financial Economics 132 (3), 76–96.

Leung, W. S., Evans, K. P., Mazouz, K., Jun. 2020. The R&D anomaly: Risk or mispricing? Journal of Banking & Finance 115, 105815.

Lin, X., Palazzo, B., Yang, F., 2020. The risks of old capital age: Asset pricing implications of technology adoption. Journal of Monetary Economics 115, 145–161.

Lopez Lira, A., 2019. Risk Factors That Matter: Textual Analysis of Risk Disclosures for the Cross-Section of Returns. SSRN Electronic Journal.

Lu, B., Ott, M., Cardie, C., Tsou, B. K., Dec. 2011. Multi-aspect Sentiment Analysis with Topic Models. In: 2011 IEEE 11th International Conference on Data Mining Workshops. pp. 81–88.

Lui, A. K. H., Lee, M. C. M., Ngai, E. W. T., Jan. 2022. Impact of artificial intelligence investment on firm value. Annals of Operations Research 308 (1), 373–388.

Manuelli, R. E., Seshadri, A., Apr. 2014. Frictionless Technology Diffusion: The Case of Tractors. American Economic Review 104 (4), 1368–1391.

Mishra, S., Ewing, M. T., Cooper, H. B., Nov. 2022. Artificial intelligence focus and firm performance. Journal of the Academy of Marketing Science 50 (6), 1176–1197.

Newey, W. K., West, K. D., 1987. A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. Econometrica 55 (3), 703–708.

Noy, S., Zhang, W., Jul. 2023. Experimental evidence on the productivity effects of generative artificial intelligence. Science 381 (6654), 187–192.

Papanikolaou, D., 2011. Investment Shocks and Asset Prices. Journal of Political Economy 119 (4), 639–685.

Pástor, L., Veronesi, P., Sep. 2009. Technological Revolutions and Stock Prices. American Economic Review 99 (4), 1451–1483.

Rock, D., May 2019. Engineering Value: The Returns to Technological Talent and Investments in Artificial Intelligence.

Seegmiller, B., Papanikolaou, D., Schmidt, L. D. W., Jan. 2023. Measuring document similarity with weighted averages of word embeddings. Explorations in Economic History 87, 101494.

Sheikh, H., Prins, C., Schrijvers, E., 2023. Artificial Intelligence: Definition and Background. In: Mission AI: The New System Technology. Springer International Publishing, Cham, pp. 15–41.

Shumway, T., 1997. The Delisting Bias in CRSP Data. The Journal of Finance 52 (1), 327–340.

Shumway, T., Warther, V. A., 1999. The Delisting Bias in CRSP's Nasdaq Data and Its Implications for the Size Effect. The Journal of Finance 54 (6), 2361–2379.

Tambe, P., 2014. Big Data Investment, Skills, and Firm Value. Management Science 60 (6), 1452–1469.

Webb, M., 2020. The Impact of Artificial Intelligence on the Labor Market.

Zhang, M. B., 2019. Labor-Technology Substitution: Implications for Asset Pricing. The Journal of Finance 74 (4), 1793–1839.

Figure 1: A Unified AI Labor and Product Indicators

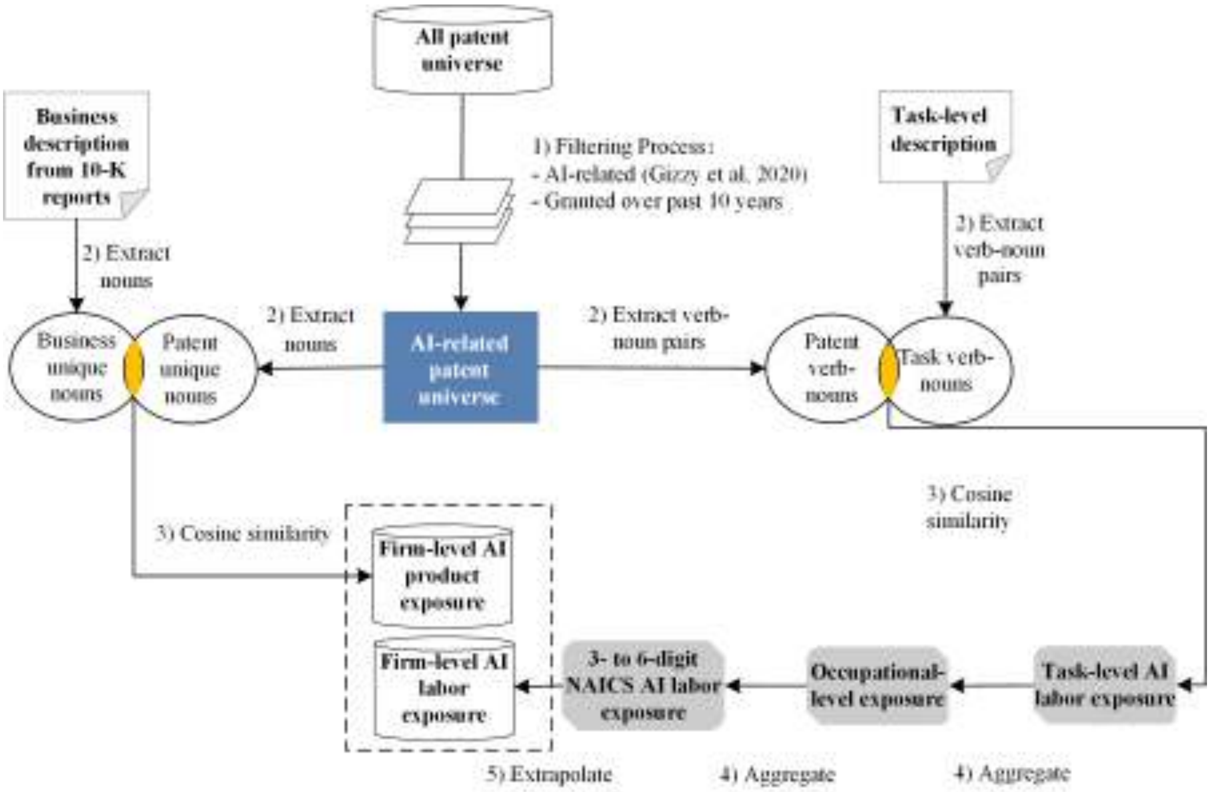


Figure 2: AI Product Exposures for Different Sectors

Note: This figure shows the heatmap of AI product exposure for different sectors. In each cell, we report the product AI exposure scaled by 100. The two-digit exposure is the average of all six-digit occupations/sectors.
The NAICS two-digit codes: Agriculture, Forestry, Fishing, and Hunting (11), Mining, Quarrying, and Oil and Gas Extraction (21), Utilities (22), Construction (23), Manufacturing (31-33, denoted as 33), Wholesale Trade (42), Retail Trade (44-45, denoted as 45), Transportation and Warehousing (48-49, denoted as 49), Information (51), Finance and Insurance (52), Real Estate and Rental and Leasing (53), Professional, Scientific, and Technical Services (54), Management of Companies and Enterprises (55), Administrative and Support and Waste Management and Remediation Services (56), Educational Services (61), Health Care and Social Assistance (62), Arts, Entertainment, and Recreation (71), Accommodation and Food Services (72), Other Services (except Public Administration) (81), and Public Administration (99).

Figure 3: AI Labor Exposures for Occupations and Sectors

Panel A: AI labor exposure for different occupations (SOC two-digit)

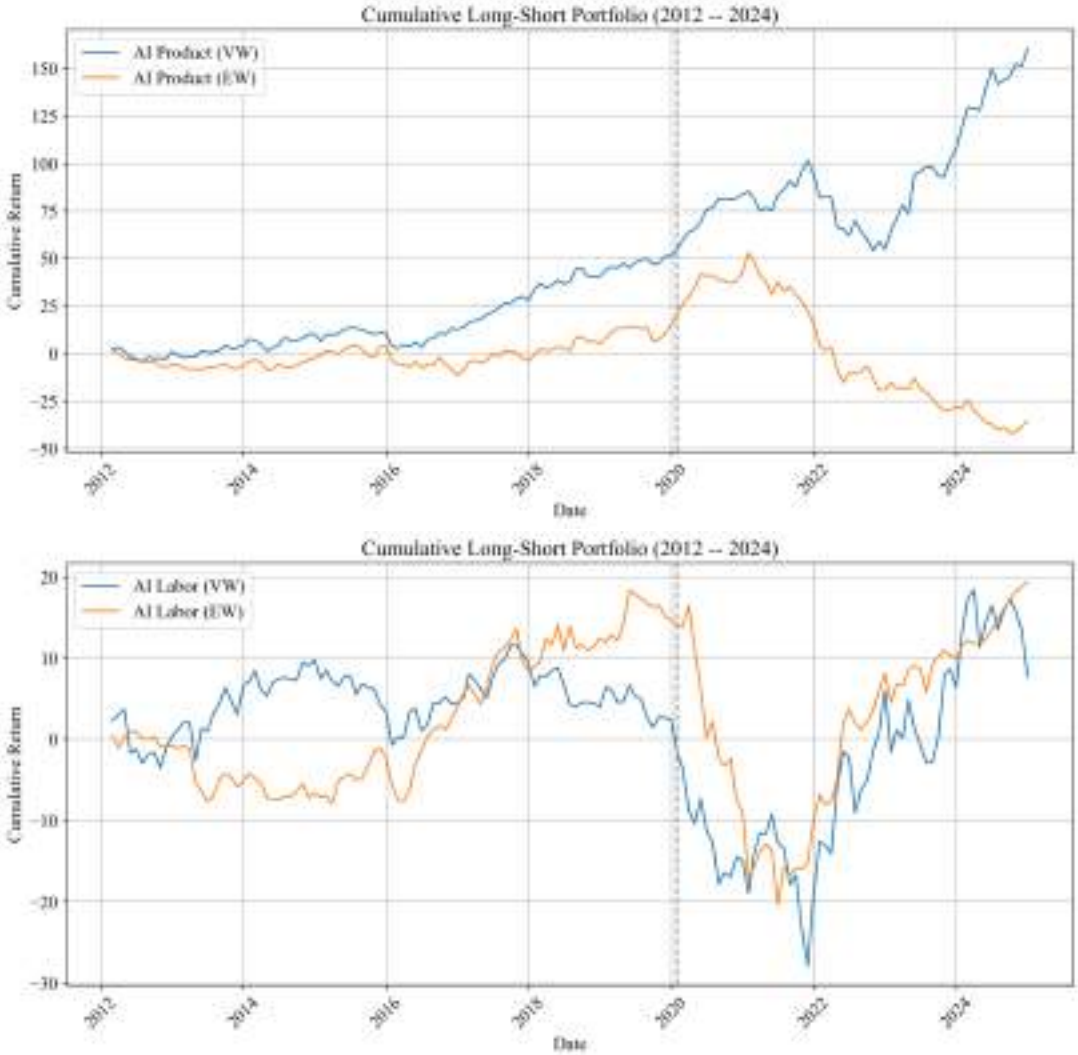
Panel B: AI labor exposure for different sectors (NAICS two-digit)

Note: This figure shows the heatmap of AI labor exposure for different occupations (Panel A) and sectors (Panel B). In each cell, we report the AI exposure multiplied by 100. The two-digit exposure is the average of all six-digit occupations/sectors.

The NAICS two-digit codes: Agriculture, Forestry, Fishing, and Hunting (11), Mining, Quarrying, and Oil and Gas Extraction (21), Utilities (22), Construction (23), Manufacturing (31-33, denoted as 33), Wholesale Trade (42), Retail Trade (44-45, denoted as 45), Transportation and Warehousing (48-49, denoted as 49), Information (51), Finance and Insurance (52), Real Estate and Rental and Leasing (53), Professional, Scientific, and Technical Services (54), Management of Companies and Enterprises (55), Administrative and Support and Waste Management and Remediation Services (56), Educational Services (61), Health Care and Social Assistance (62), Arts, Entertainment, and Recreation (71), Accommodation and Food Services (72), Other Services (except Public Administration) (81), and Public Administration (99).

The SOC two-digit codes: Management (11), Business and Financial Operations (13), Computer and Mathematical (15), Architecture and Engineering (17), Life, Physical, and Social Science (19), Community and Social Service (21), Legal (23), Educational Instruction and Library (25), Arts, Design, Entertainment, Sports, and Media (27), Healthcare Practitioners and Technical (29), Healthcare Support (31), Protective Service (33), Food Preparation and Serving Related (35), Building and Grounds Cleaning and Maintenance (37), Personal Care and Service (39), Sales and Related (41), Office and Admin Support (43), Farming, Fishing, and Forestry (45), Construction and Extraction (47), Installation, Maintenance, and Repair (49), Production (51), and Transportation and Material Moving (53).

Figure 4: Cumulative Returns of Long-Short Portfolios



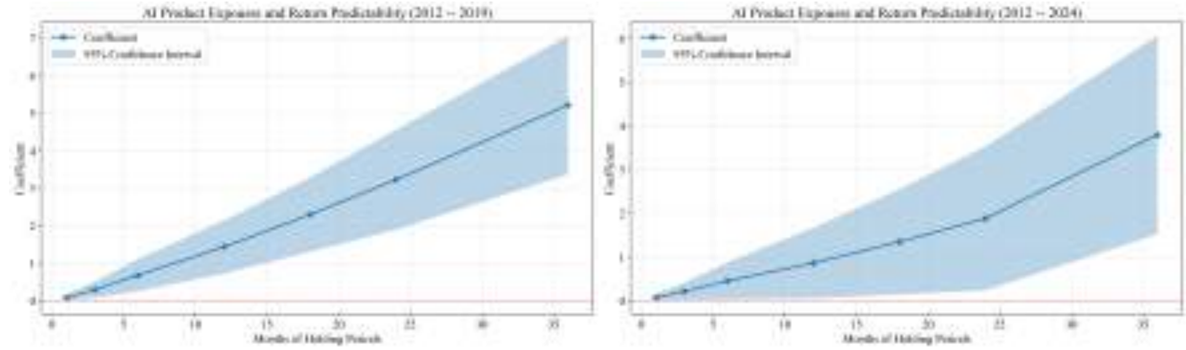
Note: This figure shows the cumulative return of the long-short portfolios. At each month, we rank companies based on their AI exposure of the previous month and take all companies below 10% and 90% quantile to form high- and low-AI portfolios, with both value-weight and equal weight weighting scheme. AI exposure are first winsorized at the 1^{st} and 99^{th} percentiles. The upper graph presents results for AI product exposure, and the lower graph for AI labor exposure. The full sample period ranges from January 2012 to December 2024.

Figure 5: Cumulative Returns of AI Premium

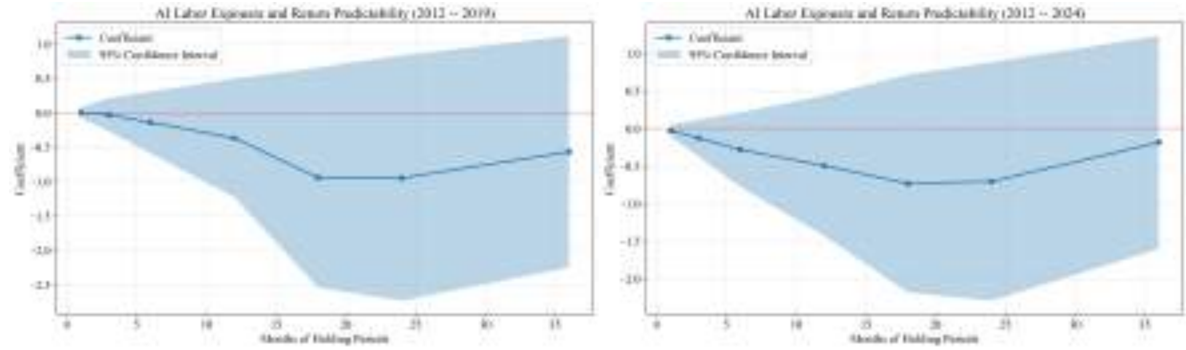
Note: This figure plots the cumulative value of the AI premium under different AI measures. To estimate the AI premium time series, at each month t, we run the cross-sectional regression with industry fixed effects and robust standard errors clustered at the firm level. We then stack β_t^{AI} over time to obtain a time series of AI premium. We consider both the AI product and labor premium. We assume that the AI measures for year m are updated six months after the fiscal year. The variables for firm characteristics are updated monthly (if there is a change). We apply the same set of control variables as in the Fama-MacBeth regression (4). A set of sector dummies based on the NAICS two-digit codes are also included. Robust standard errors clustered at firm level. All independent variables are first winsorized at the 1^{st} and 99^{th} percentiles and then standardized to zero mean and unit standard deviation. The full sample ranges from January 2012 to December 2024.

Figure 6: Predictability at Longer Periods

Panel A: AI product exposure and return predictability



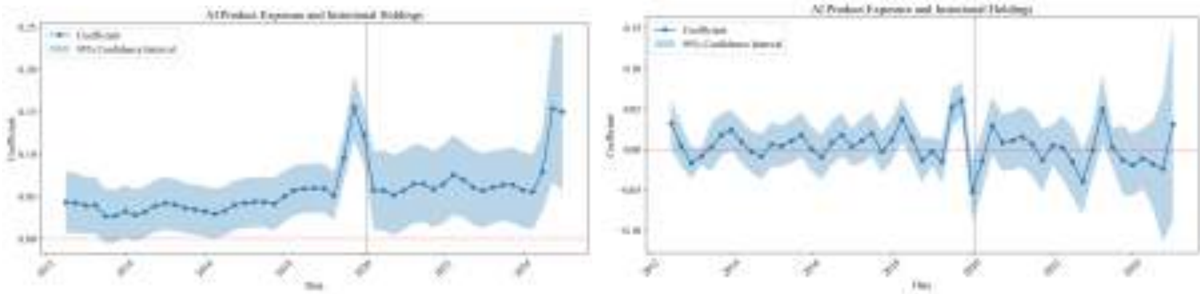
Panel B: AI labor exposure and return predictability



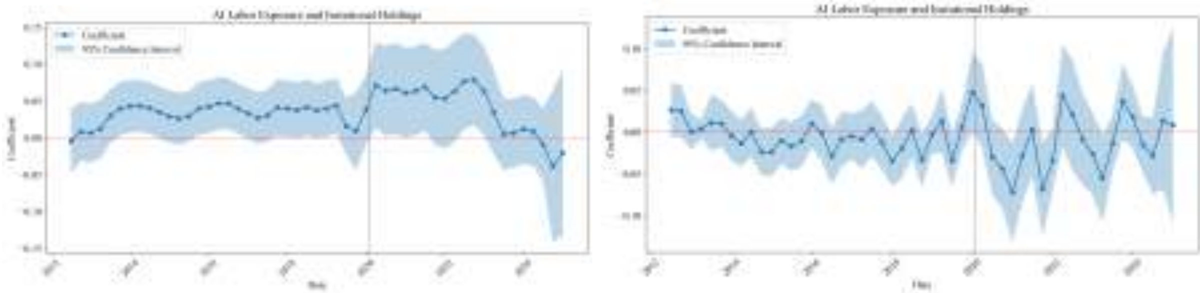
Note: This figure shows how AI exposure predicts returns over longer time horizons. The plots display coefficient estimates (dots) and 95% confidence intervals (lines) from Fama-MacBeth regressions for return horizons of 1-36 months. Panel A uses AI product exposure; Panel B uses AI labor exposure. The left-hand plots use 2012-2020 data, and the right-hand plots use 2012-2019 (pre-COVID). Other regression settings follow Table 3.

Figure 7: Institutional Holdings and AI Exposures

Panel A: Institutional holdings and AI product exposure

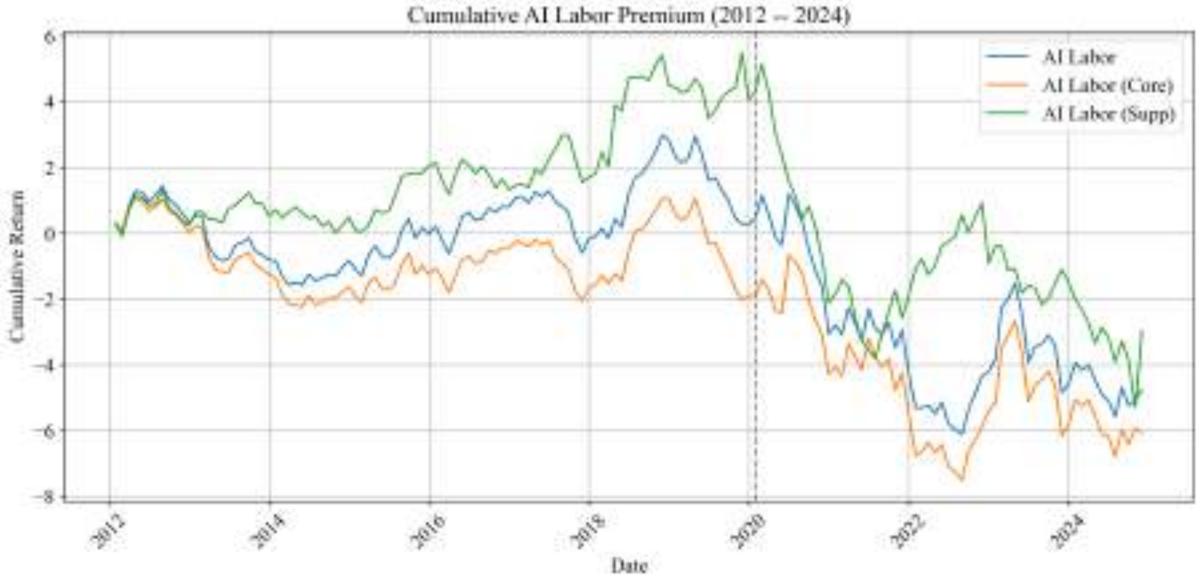


Panel B: Institutional holdings and AI labor exposure



Note: This figure shows the predictive relationship between institutional holdings and AI exposures. At each quarter-end, we run regressions of institutional holdings on AI exposures of the last quarter. We apply the same set of control variables as in the Fama-MacBeth cross-section regression (4), with robust standard errors clustered at the firm-level. Panels A and B show results for AI product exposure and labor exposure, respectively. In each panel, the left figure reports results based on the level of institutional holdings, and the right figure reports results based on quarterly changes. We plot the coefficients on the AI exposure and the corresponding 95% confidence intervals.

Figure 8: Cumulative Returns of AI Labor Premium



Note: This figure shows the cumulative value of the AI premium under different AI labor measures.We consider two types of AI labor exposure in addition to the original AI labor measure: AI^{Supp}_{Labor} that focus 100% on supplement tasks and AI^{Core}_{Labor} on core tasks. The sample ranges from January 2012 to December 2024.

Figure 9: Cumulative Value for "AI Risk" Premium

Note: This figure shows the cumulative value of the AI premium for AI risk measure. We apply the same setting as in Figure 5.

Table 1: Statistics and Correlation

Panel A: Summary statistics

	Firm-year Obs.	Mean	Std.	5%	25%	50%	75%	90%	95%
$AI_{\text{Product}} \times 100$	41596	16.449	5.453	6.968	13.149	16.693	20.145	23.270	24.972
$AI_{\text{Labor}} \times 100$	43995	0.523	0.193	0.000	0.439	0.536	0.647	0.715	0.810
AI_{Patent} (%)	13606	14.727	25.082	0.000	0.000	1.898	16.579	56.757	75.000
AI_{Employee} (%)	24043	0.119	0.289	0.000	0.000	0.000	0.097	0.347	0.646
$AI_{\text{Labor}}^{NLP} \times 100$	43995	43.667	12.097	0.000	44.872	47.623	48.804	49.922	50.533
$AI_{\text{Risk}} \times 100$	40137	12.896	4.668	5.895	9.591	12.447	15.812	19.173	21.482
$Ln(MV)$	47316	6.493	2.176	2.958	4.905	6.484	7.988	9.363	10.231
BM	47207	0.617	0.751	-0.028	0.219	0.473	0.841	1.290	1.798
OP	47276	-5.885	33.821	-68.475	-4.876	2.850	8.851	14.937	20.152
INV	46145	1.294	1.242	0.736	0.974	1.057	1.200	1.639	2.387
TAN	47063	19.597	23.169	0.370	2.497	9.970	27.340	59.334	74.401
$VOL(\%)$	47367	3.294	2.015	1.200	1.838	2.722	4.168	5.962	7.368
$LEV(\%)$	47149	23.578	23.971	0.000	3.765	17.188	36.426	55.331	69.891
$CASH$	47347	23.254	26.471	0.669	4.056	11.779	32.965	69.401	86.687
$ILLIQ$	45611	0.273	1.387	0.000	0.001	0.004	0.046	0.350	1.040

Panel B: Correlation

	AI_{Product}	AI_{Labor}	AI_{Patent}	AI_{Employee}	AI_{Labor}^{NLP}	AI_{Risk}
AI_{Product}	1					
AI_{Labor}	0.02*	1				
AI_{Patent}	0.20*	-0.01*	1			
AI_{Employee}	0.13*	0.00	0.39*	1		
AI_{Labor}^{NLP}	0.00	0.76*	-0.08*	-0.34*	1	
AI_{Risk}	0.21*	0.00	0.11*	0.10*	-0.03	1

Note: The table shows the summary statistics of all variables from fiscal year 2011 to 2023, including alternative measures of AI exposure. For the AI product exposure, Panel A provides the summary statistic of firm-level variables. $Ln(MV)$ is the natural log of market capitalization (in millions USD). BM is the book-to-market ratio. OP is the operating profit scaled by total asset of the company. INV is the asset growth, defined as the growth in the total assets for the last year ($\frac{\text{Total Asset}_{m-1}}{\text{Total Asset}_{m-2}}$). TAN is the tangibility, defined as the property, plant, and equipment divided by total assets. VOL (%) is the daily volatility calculated over the past year, $LEV(\%)$ is total debt divided by total assets. $CASH$ is the cash holding to total asset. $ILLIQ$ is the illiquidity. All variables are winsorized at level 1% and 99% quantile. Panel B shows the pair-wise correlation between different measures. "*" means significance at the level 5%.

Table 2: Portfolio Sorts on AI Exposures

Panel A: Portfolio sorts on AI product exposure level

Portfolios (%)	Mean	CAPM	FF3	FF3+UMD	FF5	FF5+UMD
Columns	(1)	(2)	(3)	(4)	(5)	(6)
H (>90%)	1.576***	0.180	0.108	0.145	0.163	0.195
	(3.51)	(0.81)	(0.54)	(0.72)	(0.83)	(0.98)
L (<10%)	0.541*	-0.422***	-0.425***	-0.431***	-0.511***	-0.506***
	(1.73)	(-2.66)	(-3.31)	(-3.31)	(-4.60)	(-4.49)
H-L (value-weight)	1.035***	0.601*	0.533*	0.576**	0.674***	0.701***
	(2.97)	(1.77)	(1.87)	(2.00)	(2.54)	(2.60)
H-L (equal-weight)	-0.231	-0.357	-0.239	-0.288	-0.116	-0.141
	(-0.84)	(-1.25)	(-1.21)	(-1.45)	(-0.72)	(-0.86)

Panel B: Portfolio sorts on of AI labor exposures

Portfolios (%)	Mean	CAPM	FF3	FF3+UMD	FF5	FF5+UMD
Columns	(1)	(2)	(3)	(4)	(5)	(6)
High AI (>90%)	1.328***	0.166	0.115	0.126	0.157	0.170
	(3.71)	(1.13)	(0.79)	(0.85)	(1.08)	(1.15)
Low-AI (<10%)	1.278***	0.076	0.026	0.079	0.062	0.111
	(3.42)	(0.47)	(0.19)	(0.58)	(0.46)	(0.82)
H-L (value-weight)	0.050	0.090	0.089	0.047	0.095	0.058
	(0.22)	(0.38)	(0.40)	(0.21)	(0.42)	(0.26)
H-L (equal-weight)	0.125	0.284*	0.227	0.194	0.241	0.195
	(0.74)	(1.69)	(1.39)	(1.18)	(1.47)	(1.19)

Note: This table presents the average monthly value-weighted return (%) of single-sorted portfolios (column 1) and its corresponding alpha in spanning regressions (columns 2-6). At the end of each month t, we sort all firms into decile portfolios by their AI exposure in month t, and we hold the both equal-weighted and value-weighted portfolios over month $t+1$. We calculate the average return, three- to six-factor alpha as well as the return performance of long-sort portfolios between the highest and lowest decile. The full sample period ranges from January 2012 to December 2024. Panel A shows the results for the AI product and Panel B the labor exposure. We provide in parentheses the t-statistics (calculated as $\sqrt{T}\bar{x}_k/sd(x_k)$, where $\bar{x}_k$ and $sd(x_k)$ are the sample average and sample standard deviation of the time series of average return for group k, respectively; T is the total number of observations in the time series). AI exposure are winsorized at the 1^{st} and 99^{th} percentiles.
* Statistical significance at the 10% level;** Statistical significance at the 5% level;*** Statistical significance at the 1% level.

Table 3: Fama-MacBeth Regressions

Dep.			R			
Periods		2012-2024			2012-2019	
Models	(1)	(2)	(3)	(4)	(5)	(6)
$AI_{Product}$	0.0742**		0.0744**	0.0801**		0.0807**
	(1.96)		(2.00)	(2.00)		(2.03)
AI_{Labor}		-0.0134	-0.0308		0.0146	0.0026
		(-0.34)	(-0.77)		(0.39)	(0.06)
$Ln(MV)$	0.0453	0.0774	0.0430	0.0420	0.0399	0.0421
	(0.44)	(0.77)	(0.41)	(0.45)	(0.39)	(0.45)
BM	0.0630	0.0831	0.0526	0.0207	0.0033	0.0170
	(0.57)	(0.94)	(0.46)	(0.23)	(0.04)	(0.19)
OP	0.7530***	0.7880***	0.7502***	0.3617***	0.4338***	0.3648***
	(3.49)	(3.53)	(3.52)	(3.88)	(4.24)	(3.92)
INV	-0.0953	-0.1348	-0.0943	-0.0253	-0.0626	-0.0247
	(-1.59)	(-2.03)	(-1.56)	(-0.54)	(-0.86)	(-0.52)
TAN	0.0398	-0.0112	0.0466	-0.1184	-0.1712	-0.1147
	(0.47)	(-0.12)	(0.53)	(-1.56)	(-1.88)	(-1.47)
VOL	-0.5052**	-0.4899**	-0.5063**	-0.4327**	-0.4174**	-0.4342**
	(-2.11)	(-2.02)	(-2.11)	(-2.89)	(-3.02)	(-2.92)
LEV	0.0336	0.0520	0.0367	0.0055	0.0197	0.0045
	(0.51)	(0.79)	(0.56)	(0.08)	(0.26)	(0.06)
$CASH$	0.2742***	0.2270	0.2775***	0.1750**	0.0462	0.1761**
	(2.72)	(1.46)	(2.74)	(2.00)	(0.27)	(2.02)
$ILLIQ$	0.0007	0.0633	0.0036	0.0496	0.1594	0.0495
	(0.01)	(0.75)	(0.07)	(1.12)	(1.52)	(1.12)
RD	0.0897	0.2082	0.0823	0.0401	0.3191	0.0403
	(1.39)	(1.28)	(1.36)	(0.69)	(1.28)	(0.68)
Observations	390573	419304	388173	247327	260991	245835
R^2	0.09	0.09	0.09	0.08	0.08	0.08
Sector Dummies	Y	Y	Y	Y	Y	Y

Note: This table reports Fama-MacBeth regressions of individual stock returns on the AI adoption measure and other firm characteristics. We first conduct the cross-sectional regressions in Equation (4) at each month t and then calculate the average of the coefficients over t. We assume that AI measure for year m is updated six months after the fiscal year-end. Firm characteristics variables are updated monthly (if there is a change). We control for market capitalization ($Ln(MV)$), book-to-market ratio (BM), operating profit rate (OP), asset growth (INV), tangibility (TAN), volatility (VOL), leverage ratio (LEV), cash holdings ($CASH$), illiquidity ($ILLIQ$) and R&D spending (RD). A set of sector dummies based on the NAICS two-digit codes are also included. At each time t, all independent variables are first winsorized at the 1^{st} and 99^{th} percentiles to reduce the impact of outliers and then standardized to zero mean and unit standard deviation. t-Statistics based on standard errors estimated using the Newey and West (1987) correction with a lag of 12 months are reported in parentheses. The R^2 is the average value of the R-squares of the cross-sectional regressions in the first step of the Fama-MacBeth procedure. The full sample period ranges from January 2012 to December 2024.
* Statistical significance at the 10% level;** Statistical significance at the 5% level;*** Statistical significance at the 1% level.

Table 4: Pooled OLS Regressions with Attention

Dep.	R					
Periods	2012-2024			2012-2019		
Models	(1)	(2)	(3)	(4)	(5)	(6)
$AI_{Product}$	0.0329	0.0438*	0.0353	0.0631**	0.1580***	0.0675**
	(1.30)	(1.72)	(1.40)	(2.30)	(3.79)	(2.45)
AI_{Labor}	0.0017	0.0004	-0.0033	0.0126	0.0093	-0.0932**
	(0.06)	(0.01)	(-0.12)	(0.43)	(0.31)	(-2.38)
$AI_{Product} \times Sent$		0.0672***			0.1509***	
		(6.82)			(3.09)	
$AI_{Labor} \times Sent$			-0.0324***			-0.1694***
			(-3.50)			(-4.37)
Observations	388173	383263	383263	245835	240925	240925
R^2	0.13	0.14	0.14	0.11	0.11	0.11
Controls	Y	Y	Y	Y	Y	Y
Month Dummies	Y	Y	Y	Y	Y	Y
Sector Dummies	Y	Y	Y	Y	Y	Y

Note: This table reports pooled OLS regressions of individual stock returns on the AI adoption measure and AI attention. We assume that AI measure for year m is updated six months after the fiscal year-end. Firm characteristics variables are updated monthly (if there is a change). We apply the same set of control as in the Fama/MacBeth regression. A set of sector dummies based on the NAICS two-digit codes are also included. At each time t, all independent variables are first winsorized at the 1^{st} and 99^{th} percentiles to reduce the impact of outliers and then standardized to zero mean and unit standard deviation. t-Statistics based on robust standard errors clustered at the firm-level are reported in parentheses. * Statistical significance at the 10% level;** Statistical significance at the 5% level;*** Statistical significance at the 1% level.

Table 5: Portfolio Sorts with Different AI Labor Exposures (Pre-COVID)

		Mean	CAPM	FF3	FF3+UMD	FF5	FF5+UMD
AI_{Labor}	H-L(VW)	0.026	-0.096	-0.017	0.040	0.026	0.106
		(0.15)	(-0.52)	(-0.09)	(0.22)	(0.15)	(0.64)
	H-L(EW)	0.155	0.180	0.217	0.202	0.252	0.248
		(1.01)	(1.10)	(1.31)	(1.19)	(1.57)	(1.50)
AI^{Supp}_{Labor}	H-L(VW)	-0.194	-0.181	-0.059	-0.043	-0.076	-0.056
		(-1.02)	(-0.89)	(-0.32)	(-0.23)	(-0.41)	(-0.30)
	H-L(EW)	0.155	0.210	0.258	0.210	0.296*	0.255
		(0.97)	(1.24)	(1.52)	(1.22)	(1.78)	(1.50)
AI^{Core}_{Labor}	H-L(VW)	-0.119	-0.272	-0.179	-0.139	-0.145	-0.088
		(-0.68)	(-1.51)	(-1.04)	(-0.79)	(-0.89)	(-0.53)
	H-L(EW)	0.030	-0.020	0.037	0.038	0.072	0.083
		(0.19)	(-0.12)	(0.22)	(0.22)	(0.44)	(0.50)

Note: This table presents the average monthly returns (%) of long-short portfolios sorting on different measure and their corresponding alphas from spanning regressions. We consider two types of AI labor exposure in addition to the original AI labor measure: AI^{Supp}_{Labor} that weights 100% on supplement tasks and AI^{Core}_{Labor} on core tasks. Each month, we rank firms based on their AI exposure from the previous month and form high- and low-AI portfolios using companies below the 10th and above the 90th percentiles, employing both value-weighted and equal-weighted schemes. The sample period spans from February 2012 to December 2019. Panel A reports results for AI product exposure, and Panel B for labor exposure. We report t-statistics in parentheses, calculated as $\sqrt{T} \cdot \bar{x}_k/\text{sd}(x_k)$, where $\bar{x}_k$ and $\text{sd}(x_k)$ are the sample mean and standard deviation of the time series of average returns for group k, respectively, and T is the total number of time series observations. AI exposures are winsorized at the 1st and 99th percentiles. The sample period ranges from January 2012 to December 2019.
* Statistical significance at the 10% level;** Statistical significance at the 5% level;*** Statistical significance at the 1% level.

Table 6: AI Labor and Company Size During the Pandemic

Dep.	R					
AI Labor Expousre	AI_{Labor}		AI_{Labor}^{Core}		AI_{Labor}^{Supp}	
Models	(1)	(2)	(3)	(4)	(5)	(6)
AI_{Labor}	-0.3426	-0.5883*	-0.2441	-0.454	-0.6773***	-0.8346***
	(-1.22)	(-1.82)	(-0.92)	(-1.52)	(-2.74)	(-3.02)
$AI_{Labor} \times MV$		0.6387**		0.5907**		0.6547**
		(2.43)		(2.33)		(2.42)
Observations	33343	33343	33343	33343	33343	33343
R^2	0.14	0.14	0.14	0.14	0.14	0.14
Controls	Y	Y	Y	Y	Y	Y
Month Dummies	Y	Y	Y	Y	Y	Y
Sector Dummies	Y	Y	Y	Y	Y	Y

Note: This table reports results from pooled OLS regressions of monthly stock returns on AI labor exposure and its interaction with firm size. The sample period is January to December 2020. All specifications include controls for firm characteristics, as well as month and sector fixed effects. Robust standard errors are clustered at the firm-level.
* Statistical significance at the 10% level;** Statistical significance at the 5% level;*** Statistical significance at the 1% level.

Table 7: Fama-MacBeth Regressions with AI Risk Exposure

Dep.					R				
Periods		2012-2024			2012-2019			2023-2024 (Pooled OLS)	
Models	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
AI_{Risk}	0.0862**	0.0759**	0.0793**	0.0911**	0.0832**	0.0848**	0.0368	0.0407	0.0576
	(2.56)	(2.39)	(2.48)	(2.48)	(2.44)	(2.41)	(0.46)	(0.51)	(0.70)
$AI_{Product}$		0.0671*	0.0594*		0.0749*	0.0727*		0.0119	-0.0135
		(1.82)	(1.65)		(1.82)	(1.84)		(0.14)	(-0.17)
$AI_{Product} \times AI_{Risk}$			0.0368			0.0091			0.1377*
			(1.40)			(0.44)			(1.76)
AI_{Labor}		-0.0274	0.0368		0.0112	0.0091		-0.1769	0.1377
		(-0.67)	(-0.66)		(0.27)	(0.26)		(-2.28)	(-2.26)
Observations	378556	378556	378556	239766	239766	239766	46193	46193	46193
R^2	0.09	0.10	0.10	0.08	0.08	0.08	0.16	0.16	0.16
Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y
Sector Dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y

Note: This table reports Fama-MacBeth regressions of individual stock returns on AI risk exposure and its interaction with AI product exposure. We consider three sub-periods: 2012-2024 (columns 1-3), 2012-2020 (columns 4-6) and 2023-2024 (columns 7-9). For columns 1-6 we apply the same Fama-MacBeth cross-section setting as in Table 3 and for columns 7-9, we use pooled OLS with sector-month dummies and robust errors clustered at the firm-level.
* Statistical significance at the 10% level;** Statistical significance at the 5% level;*** Statistical significance at the 1% level.

A Variable Definitions

Table G.8 shows the definition of firm characteristics used in our analysis.

B Constructing the AI Product Exposure from 10-K Filings

Procedure

We integrate both the AI patent database and the firm's 10-K filings to develop the AI product measure. We determine the similarity score between each firm's business description and the abstracts of AI patents by following the method described by Hoberg and Phillips (2016). In their study, they calculated the similarity scores of business descriptions across various companies to establish industry classifications. This similarity score is derived from the level of overlap between the unique nouns of two firms. Similarly, we calculate the similarity score between the business description and AI patent abstracts. A company with significant AI product exposure is expected to have its business "spanned" by the AI patent universe. Consequently, the similarity score between the company's business description and AI patents should be high.

Specifically, we first perform regex searching to isolate the Item 1 text, where we specify the start and end of the Item 1 Section.[13] Next, we extract unique nouns from all 10-K filings of that year to obtain the unique noun universe. We follow Hoberg and Phillips (2016) to remove nouns that appear in more than 25% of all companies and proper nouns related to countries and cities. On average, there are 20,000 unique nouns in our sample extracted from the 7,000 business descriptions each year. In the second step, we extract the nouns from all patents from the past 10 years. We define unique nouns as those that appear in more than 1% but less than 25% of the total patent universe. On average, we have around 200 nouns extracted from more than 200,000 patents each year. Table G.9 presents the top 25 nouns with the highest frequency. More than 20% of patents mentioned information. Interestingly, the "intelligence" from AI does not appear (though there are 'ai' in some instances).

Second, for each year m, we first construct the business descriptions noun universe by extracting all unique nouns from that year's business descriptions. Suppose this business descriptions noun universe in year m contains W words (e.g., 20,000 unique nouns). We then extract the unique nouns from the patent universe in year m and map these (e.g., 600 unique nouns) into the business descriptions noun universe, obtaining a $W \times 1$ indicator vector $V_{\text{Patent},m}$, whose entry equals one if the corresponding noun in the business

[13] The regex we are using is: "((?:ITEM|Item)" \s+1(?![0-9])(?![A-Za-z])[.:\s—]* \s*[\w\W]*?)(?=(?:ITEM|Item)\s+1A?̇|(?:ITEM|Item)\s+2(?![A-Za-z])|PART\s+II(?![A-Za-z]))".

descriptions universe appears in the patent noun universe and zero otherwise.

Similarly, for each company i, we extract the nouns from its business description in the year m (e.g., 40 nouns for a typical firm) and map them into the noun universe of all business descriptions in the year m, yielding another $W \times 1$ indicator vector $V_{i,m}$, whose entry equals one if the corresponding noun in the entire business description universe appears in company i's business description noun universe and zero otherwise. We normalize both vectors to have unit length. The AI product exposure of company i in year m is then defined as the cosine similarity between the company vector and the patent vector:

$$\text{AI}_{i,m}^{\text{Prod}} = V_{i,m}^{\top} V_{\text{Patent},m}.$$

Examples of Companies with High AI Product Exposure

In this subsection, we present examples of non-tech companies with relatively high exposure to AI products. We extract phrases related to AI from their business descriptions for the years 2009, 2015, and 2023, respectively.

Education Sector: Stride, Inc. (previously K12 Inc.)

- CIK number:1157408; AI product exposure: 0.23 (2009), 0.24 (2015), 0.17 (2023).
- Company Overview (2023): "We are an education services company providing virtual and blended learning. Our technology-based products and services enable our clients to attract, enroll, educate, track progress, and support students. These products and services, spanning curriculum, systems, instruction, and support services, are designed to help learners of all ages reach their full potential through inspired teaching and personalized learning. Our clients are primarily public and private schools, school districts, and charter boards. Additionally, we offer solutions to employers, government agencies, and consumers."
- Excerpts from business description (2009)
    1. Data processing: "Student Administration Management System (SAMS) is our proprietary Student Information System. SAMS is integrated with the OLS and several other proprietary systems, including our Online Enrollment System that allows parents to complete school enrollment forms online, and our Order Management System (OMS), which generates orders for offline learning kits and computers to be delivered to students."
    2. Data processing: "Our progress tracking tool allows students, parents, and teachers to monitor student progress. In addition, **information collected**

by our progress tracking tool regarding student performance, attendance, and other data is transferred to our proprietary management system for use in providing administrative support services."

- Excerpts from business description (2015)
  1. Service customization: " Our instructional system allows students to learn from a curriculum that **caters to their unique learning style and offers a high degree of program flexibility**. Certain adaptive learning features integrated into some curricular products can individualize lessons based on the level of student comprehension."
  2. Service customization: "Adaptive Learning: We have **learning management systems** and can now build courses that are adaptive, which enables individualized learning experiences as the course 'adapts' at key points to student behavior and input."
- Excerpts from business description (2023)
  1. Service customization: "This includes pursuing the development and licensing of curricula and platforms that are accessible from tablets and mobile devices, as well as **leveraging adaptive learning technologies and solutions.** We will also invest in our current products and assets to make them more accessible to larger markets by improving the user experience and content."
  2. Data Analysis: "Our end-to-end platform includes content management, learning management, **student information, data reporting and analytics,** and various support systems that allow customers to provide a high-quality and personalized educational experience for students."

Arts and Entertainment: Everi Holdings Inc. (Previously, Global Cash Access Holdings, Inc.)

- CIK number: 1318568; AI product exposure: 0.19 (2009), 0.22 (2015), 0.22 (2022).
- Company Overview: "Everi develops and offers products and services that provide gaming entertainment, improve our customers' patron engagement, and help our casino customers operate their businesses more efficiently. We develop and supply entertaining game content, gaming machines, gaming systems, and services for land-based and iGaming operators. Everi is a provider of financial technology solutions that power casino floors, improve operational efficiencies, and fulfill regulatory requirements. The Company also develops and supplies player loyalty tools and mobile-first applications that enhance patron engagement for our customers and venues in the casino, sports, entertainment, and hospitality industries."

- Excerpts from business description (2009)

  1. Data analysis: “QuikReports is a browser-based reporting tool that provides marketing professionals with real-time access to and **analysis of information** on patron cash access activity.””
  2. Data analysis: “Central Credit is the leading gaming patron credit bureau that allows **gaming establishments to improve their credit-granting decisions.** Our Central Credit database contains decades of gaming patron credit history and transaction data on millions of gaming patrons.”

- Excerpts from business description (2015)

  1. Data analysis: “Database services that allow gaming establishments access to information from our proprietary patron transaction database for the purposes of player acquisition, direct marketing, **market share analysis,** and a variety of other patron promotional uses. Our proprietary patron transaction database includes information that is captured from the transactions we process.””
  2. Data processing: “We conduct research and development activities primarily to develop gaming systems, gaming engines, casino **data management systems**, casino central monitoring systems, video lottery outcome determination systems, gaming platforms, and gaming content, as well as to add enhancements to our existing product lines.”
  3. Data analysis: “Products and services that improve **credit decision making,** automate cashier operations, and enhance patron marketing activities for gaming establishments.”

- Excerpts from business description (2023)

  1. Data analysis: “Central Credit is our gaming patron credit bureau service, which, on a subscription basis, allows gaming operators to improve their **credit-granting decisions by obtaining access to a database containing credit information and transaction data on millions of gaming patrons**.”
  2. Fraud analysis: “Everi Compliance is a leading Anti-money Laundering (“AML”) **management tool** for the gaming industry.”

Healthcare: Laboratory Corporation of America Holdings

- CIK number: 920148; AI product exposure: 0.24 (2009), 0.26 (2015), 0.24 (2023).

- Company Overview: “Laboratory Corporation of America Holdings and its subsidiaries (the ’Company’), headquartered in Burlington, North Carolina, is the second largest independent clinical laboratory company in the United States based on 2014 net revenues. Through its national network of laboratories, the Company offers a broad range of clinical laboratory tests that are used by the medical profession in core testing, patient diagnosis, and in the monitoring and treatment of diseases.”
- Excerpts from the business description (2009)
    1. Process automation: “The Company provides complete viral load testing, as well as HIV genotyping and phenotyping. In 2000, the Company added HIV GenoSure™ to its portfolio of HIV resistance testing services. The Company’s use of this leading-edge technology put it in the forefront of HIV drug resistance testing, one of the most important issues surrounding the treatment of HIV. In 2007, the Company became the first commercial laboratory to offer fully **automated** real-time HIV testing from Roche Diagnostics. Additionally, the Company provides comprehensive testing for HCV including both PCR testing and genotyping at CMBP, NGI and Viro-Med.”
    2. Data analysis: “the Company’s connectivity platform integrates easily with a wide variety of existing **electronic medical record systems**, practice management systems, and procedure writing systems... “
- Excerpts from business description (2015)
    1. Data analysis: “LCD has new population health analytics programs in development to provide **healthcare business intelligence tools** to hospitals, physician practices, and accountable care organizations (ACOs). These tools are intended to assist customers in their compliance and reporting requirements with respect to the efficient management of their productivity, quality, and patient outcome metrics.”
    2. Data analysis: “Xcellerate Insights supports the analysis of a trial’s most recent operational metrics in a secure collaborative portal. In addition to Xcellerate, CDD’s proprietary technology assets include CDD’s investigator database and **analytic methodologies** utilized to design and manage patient enrollment, site selection, and investigator selection to produce higher quality and faster clinical trials.”
    3. Monitoring services: “CDD and LCD are also collaborating to use LCD information to support clinical **trial recruitment and post-trial monitoring.**”
- Excerpts from business description (2023)

1. For several years, the Company has deployed **artificial intelligence and machine learning tools (AI)** to supplement its existing data analysis projects and support greater efficiency in its operations. The Company is currently evaluating and testing various additional applications of AI while also implementing policies and processes to provide appropriate governance over the use of AI by the Company."

Financial: VISA INC.

- CIK number: 1403161; AI product exposure: 0.22 (2009), 0.16 (2015), 0.19 (2023).
- Company Overview (2023): "Visa is one of the world's leaders in digital payments. Our purpose is to uplift everyone, everywhere, by being the best way to pay and be paid. We facilitate global commerce and money movement across more than 200 countries and territories among a global set of consumers, merchants, financial institutions, and government entities through innovative technologies."
- Excerpts from business description (2009)
  1. Fraud detection: "Visa Advanced Authorization includes **preventive, monitoring, investigative, and predictive** tools that are intended to mitigate and help eliminate fraud at the cardholder and merchant levels."
  2. Fraud detection: "Our fraud detection and prevention offerings include Verified by Visa, a global internet authentication product that permits cardholders to authenticate themselves to their issuing financial institution using a unique personal code; Visa Advanced Authorization adds additional fraud detection capability by providing **real-time risk scores to authorization messages**."
- Excerpts from business description (2015)
  1. Fraud detection: "Fraud and **data analytics**: As an industry leader in payment security, we have enhanced our real-time data analytics capabilities. When combined with Visa's centralized network structure, these capabilities help our financial institution clients and merchants identify and address fraud."
- Excerpts from business description (2023)
  1. Risk monitoring: "The integration of technology, like generative AI, can create new and better offerings that compete with our value added services, such as strengthened **risk monitorization and managing digital identification**.""
  2. "Solutions such as Visa Advanced Authorization, Visa Secure, Visa Risk Manager, Decision Manager, Visa Consumer Authentication Service, and payment-decisioning solutions from CardinalCommerce empower financial institutions

and merchants with tools that help automate and simplify fraud prevention and enhance payment security."

3. "Visa's risk and identity solutions transform data into insights for near real-time decisions and facilitate account holder authentication to help clients prevent fraud and protect account holder data."

C Constructing the AI Labor Exposure

A large part of the literature on measuring AI labor exposure is based on the logic of "AI technology – task/ability – occupation – firm/industry". The difference lies in how to define the connection between tasks/abilities and AI technology. The first strand of studies uses crowd-sourcing services like Amazon Mechanical Turk or CrowdFlower (e.g., Felten et al., 2019, Felten et al., 2018 and Brynjolfsson et al., 2018). With the advance of ChatGPT, another strand of studies uses LLM models to determine the task-level exposure of AI technologies (e.g., Eloundou et al., 2024 and Eisfeldt et al., 2025), though using ChatGPT prompts may suffer from replicability issues. Finally, Webb (2020) uses AI patent information to determine the exposure of tasks to AI technology. For the main analysis, we adopt the verb-noun mapping method from Webb (2020), with several key changes and improvements to the procedure.

Step 1: Determining Task-level AI Exposure

The task-level AI exposure refers to the extent to which a task can be automated using AI technology. Addressing this concept involves three significant issues. First, it is necessary to define the specific AI technologies influencing labor. Second, tasks must be accurately defined within each occupation. Third, a method for measuring the relationship between AI technology and tasks should be established.

For the first issue, patent information is considered a useful resource. The advancements in AI technology are often reflected in AI-related patents. We use datasets from the US Patent and Trademark Office (USPTO), covering all patents filed in the US since 1982.[14] To classify patents as AI or non-AI, we employ the classification method developed by Giczy et al. (2021), which combines both machine learning and manual verification to identify AI-related patents. Our analysis includes all granted patents

For the second issue, we utilize the O*NET database, which provides information on over 900 occupations, including their respective Detailed Work Activities (DWAs) and tasks. A DWA encompasses a comprehensive action as part of completing a task, whereas a task represents an occupation-specific unit of work potentially linked to zero,

[14]https://www.uspto.gov/ip-policy/economic-research/research-datasets/artificial-intelligence-patent-dataset

one, or multiple DWAs. Our analysis employs the task descriptions. As we estimate labor exposure from 2013 to 2024, we also incorporate different versions of the O*NET database, depending on the year of evaluation.

Finally, to measure the AI technology–task connection, we follow the idea of Webb (2020) to use the verb-noun matching strategy but with a different setting. In Webb (2020)'s setting, for each year m, we extract verb-noun pairs from the AI-related patent abstracts, which we denote as the "verb-noun universe". Then, for each task description, we again extract verb-noun pairs (assuming we have K_i verb-noun pairs in each task). For each verb-noun pair k extracted from the task description, we count its occurrence (as a percentage, $p_{k,m} = \frac{\text{Occurrence}_{k,m}}{\text{Total Pairs}_m}$) in the verb-noun universe, and a higher level of percentage signifies greater exposure of the task to AI technology.

We employ cosine similarity, a method commonly utilized in semantic analysis, to measure the relatedness between tasks and technologies. Initially, we construct two distinct verb-noun universes: one derived from all tasks described in O*NET for a given year, and the other from all relevant patents filed during that same year. Following Honnibal and Johnson (2015), we use the "spaCy" library in Python to extract verb-noun pairs from the text. For the patent verb-noun universe, we retain only those pairs that appear in more than 0.1% but no more than 25% of the documents in that universe. This filtering process aims to isolate verb-noun pairs representative of the AI-related knowledge and capabilities of the given year. At year m, we consider all patents from the preceding 10 years, including year m itself, resulting in an average of over 250,000 AI-related patents within each window. Each year, we identify approximately 600 verb-noun pairs from the patent universe and around 19,000 pairs from the task universe.

We emphasize several critical considerations in handling the patent and occupational data. First, while capturing verb-noun pairs, we deliberately avoided consolidating nouns from child groups into parent groups (e.g., change "bread" to "food",as in Webb, 2020). This decision reflects the expectation that both task descriptions and patent abstracts are written in highly concise and precise language. For instance, a bakery does not deal with meat, so patents concerning meat processing are unlikely to apply to a bakery. Although this approach may significantly reduce the size of the cosine similarity, it enhances the accuracy of our measurement.

Secondly, while our paper focuses on labor exposure from 2012 to 2024, our estimation of task exposure considers patents over the preceding decade up to the estimation year. For instance, tasks in 2014 are exposed to patents from 2005 to 2014, whereas tasks in 2020 are exposed to all AI-related patents from 2011 to 2020. This approach aims to balance the evolving technological landscape – given that AI technology changes annually and older technologies become outdated – while accurately reflecting the current technology available to companies.

Table G.10 presents examples of verb-noun pairs extracted from tasks. In patent text, our method successfully captures functions related to the AI patent, particularly those involving data handling. In task text, the verb-noun pairs reflect both the action (e.g., 'coordinate', 'activity') and the goal of the task (e.g., 'increase', 'efficiency').

Table G.11 presents the top 25 verb-noun pairs for 2009, 2015, and 2023, extracted from the patent universe, ranked by frequency. Tasks with the highest exposure predominantly involve data analysis and data processing.

Next, we construct 0-1 vectors between the "task" and "patent" verb-noun universes. Assuming we have W words in the "task" verb-noun universe, the 0-1 vector is a W-by-1 vector where each position equals one if the corresponding verb-noun in the "task" is found in the "patent" verb-noun universe ($U_{\text{Patent},m}$). Similarly, we also construct a W-by-1 vector for each task. We then normalize each vector to have unit length.

For each task, we calculate the cosine similarity between the verb-noun and use it as the task-level exposure :

$$\text{AI}^{\text{Task}}_{k,m} = U^{\top}_{k,m} U_{\text{Patent},m} \cdot, \tag{10}$$

for task i in year m. The cosine similarity is better than counting the frequency in that it does not depend on the length of the task description. Table G.12 shows the top 20 tasks with the highest exposure over the sample period. The tasks with high AI exposure are either those involving repeated actions (e.g., "Observe workers operating equipment") or those related to data analysis (e.g., "Analyze clinical or survey data").

Step 2: Estimating Occupational and Industry-level AI Exposure

Once we have the task-level exposure, the occupation-level AI exposure is constructed as

$$\text{Occupation-level AI Exposure}_{j,m} = \frac{1}{N_m} \sum_{i=1}^{N_m} \text{Task-level AI Exposure}_{i,m} \times \text{Task Relevance}_{i,m}, \tag{11}$$

where Occupation-level AI Exposure$_{j,m}$ is the occupation-level AI exposure for occupation j in year m, Task-level AI Exposure$_{i,m}$ is the task-level AI exposure for task i in year m, and N_m is the total number of tasks for occupation j in year m. Task Relevance$_{i,m}$ is the relevance weight of each task to the occupation, which takes three values: "Core", "Supplemental" or "Not Available". The relevance is estimated by job incumbents who rated the provided task as relevant to their job.[15] Task Relevance$_{i,m}$ is set to 0.7 for core tasks and 0.3 for supplemental tasks; for tasks without a label, we set the value at 0.5. The portion of tasks without importance classification is less than 5% of the total number of tasks.

Finally, we construct the AI exposure for the 6-digit NAICS industry based on the

[15] see: https://www.onetonline.org/help/online/scales

Bureau of Labor Statistics (BLS) industry occupation data[16]

$$\text{AI}_{s,m}^{\text{Labor}} = \sum_{j=1}^{M_m} \text{Occupation-level AI Exposure}_{j,m} \times \text{Weight}_{j,m}, \tag{12}$$

where $\text{AI}_{s,m}^{\text{Labor}}$ is the industry-level AI for industry s in year m and M_m is the total number of occupations in industry s. $\text{Weight}_{j,m}$ is the employment percentage of occupation j in industry s. $\text{Weight}_{j,m}$ is not available for the years 2013, 2015, and 2017 during our sample period because the BLS conducted the survey to collect industry-occupation employment data every two years during the early years. For missing year data, we interpolate between adjacent years. For example, the employment data in a certain industry for a certain occupation in 2013 is the average between 2012 and 2014.

Under the above procedure, the industry-level AI labor exposure is time-varying from the following aspects:

- Changing AI technology: the landscape of the AI patent universe is changing each year, which means that the task-level exposure in Eq. (10) is varying as the verb-noun pair coverage is time-varying.
- Changing occupation definition: From 2012 to 2023, there are emerging and changing definitions of occupations. Following the BLS industry-occupation employment data, we consider different versions of ONET data across our sample period. For 2012–2015, we use ONET 18.0; for 2016–2017, version 22.0; for 2018, 23.0; for 2019, 24.0; for 2020, 25.1; and for 2023, 29.0.
- Changing industry-occupation distribution: the occupational mix in each industry changes every year.

Finally, Table G.13 presents the top five occupations in terms of average AI exposure over the sample period. Occupations with the highest AI exposure are those involving substantial data and information processing, such as market data analysts and property appraisers. In contrast, occupations with the lowest AI exposure are those requiring interpersonal interactions (e.g., singers or labor relations specialists) and service-related tasks (e.g., maids). These results are consistent with the findings in Table G.11, where most verb–noun pairs derived from AI patents relate to data and information processing.

[16]: https://www.bls.gov/emp/data/projections-archive.htm

D Robustness: Using Text Embeddings to Evaluate Task Labor Exposure

While the verb-noun matching strategy may effectively capture how tasks are affected by AI technology, as represented by the patent universe, it fails to incorporate contextual information. This limitation may lead to an underestimation of exposure. As shown in Table 1, the magnitude of labor exposure is significantly smaller than that of product exposure, indicating fewer verb-noun matches.

To address this issue, we employ a state-of-the-art text-embedding technique. This approach encodes a document's semantic meaning into a geometric vector representation. Under the same embedding model, documents with similar structures will exhibit high vector similarity.

We employed the "QWEN3-Embedding-8B" model, developed by Alibaba DAMO Academy, which comprises 8 billion parameters. As of September 2025, when we prepared the results, this model ranked first in several text analysis tasks.[17] With its larger parameter size, the model converts text into longer vectors compared to older models – specifically, 4096-dimensional embedding vectors, whereas smaller models typically output 1024-dimensional vectors. This capacity enhances the model's ability to capture semantic information. Subsequently, we calculated task-level AI exposure by determining the cosine similarity between the text embeddings of the task description and the patent abstract:

$$\text{Task-level AI Exposure}_i \equiv Simi_i = \frac{1}{J}\sum_{j}^{J}(V_i \cdot V_j) \tag{13}$$

for task i, $j = 1, ..., J$, where J is the total number of patents in the universe. In the above definition, the AI exposure for task i is the average of its similarity with all patents, with the value ranging from 0 to 1. After obtaining the task-level exposure, we apply the same procedure as in the verb-noun case.

E Getting Firm-level Institutional Ownership Data from 13-F Reports

The 13-F reports are quarterly documents detailing the holdings of investment managers or institutions managing assets exceeding one million USD. Each investment manager must disclose their long positions in stocks and stock options, specifying both the market value and the total number of shares. This information allows researchers to ascertain the proportion of a company's shares held by institutional investors.

[17] huggingface.co/Qwen/Qwen3-Embedding-8B

To gather the required information, we initially downloaded all reports from 2012 to 2024 using the SEC API. Subsequently, for each company – identified and located via the CUSIP code – we aggregated the number of shares reported by all managers. Finally, we calculated our institutional ownership measure by

$$IO_{i,n} = \frac{\text{Shares Hold by Institutional Investors}_{i,n}}{\text{Total Shares Outstanding}_{i,n}}, \tag{14}$$

for each company i in each quarter n.

F Definition of Risk Types

Below is the full list of definition of risk topics and some examples, which refers to Bao and Datta (2014):

- "Financial condition risks" are factors related to a history of loss, resulting in poor financial conditions. For example: "We have experienced net losses, and we may not be profitable in the future."
- "Restructuring risks" involve a situation where the target company has filed for bankruptcy protection, or the company mentioned it is undergoing restructuring. For example: "We may need to incur impairment and other restructuring charges, which could materially affect our results of operations and financial conditions."
- "Funding risks" refer to the inability to raise capital to expand, for normal operations, or to match competition. For example: "Banctrust may need to raise capital in the future when capital may not be available on favorable terms or at all."
- "Merger & Acquisition risks" encompass any factor that is related to M&A, such as acquisitions not meeting expectations, or high M&A costs. For example: "Implementing our acquisition strategy involves risks, and our failure to successfully implement this strategy could have a material adverse effect on our business."
- "Regulation changes" are risks about government regulation changes, including environmental, accounting, or privacy laws. For example: "The company is subject to environmental regulations and liabilities that could weaken operating results."
- "Catastrophes" include natural disasters or terrorist attacks. For example: "Future terrorist attacks may have a material adverse impact on our business."
- "Shareholder's interest risks" include situations where: (1) the holder's interest is different from the shareholders, (2) the shareholder has very strong control power (few large shareholders), or (3) there is no large shareholder. For example: "We may encounter conflicts of interest with our controlling stockholder."

- “Macroeconomic risks” include economic downturn, financial crisis, high energy prices, inflation, or recession. For example: “Demand for our products will be affected by general economic conditions.”
- “International risks” are factors related to global operations, including currency and exchange rate risks. For example: “Our international operations are subject to many uncertainties, and a significant reduction in international sales of our products could adversely affect us.”
- “Intellectual property risks” involve the possibility that the target company may infringe or be infringed by another company’s patents. For example: “We may not be successful in adequately protecting our intellectual property.”
- “Potential defects” in products refer to product liabilities or any risks related to product defects. For example: “We may incur substantial costs as a result of warranty and product liability claims which could negatively affect our profitability.”
- “Potential/Ongoing Lawsuits” are current or ongoing significant litigation or lawsuits. For example: “We are currently subject to securities class action litigation, the unfavorable outcome of which might have a material adverse effect on our financial condition, results of operations and cash flows.”
- “Infrastructure risks” are related to changes, upgrades, or maintenance of the target company’s infrastructure, which includes distribution networks, IT, or organizational infrastructure. For example: “The infrastructure of our transmission and distribution system may not operate as expected, and could require additional unplanned expense which would adversely affect our earnings.”
- “Disruption of operations” are risks that operations may be disrupted due to a complex manufacturing process or software systems. For example: “Material disruption to our manufacturing plants in Wisconsin could adversely affect our ability to generate revenue.”
- “Human resource risks” involve attracting, recruiting, and maintaining key personnel or employees, such as the CEO, executives, R&D staff, or sales people. For example: “We depend upon our key personnel and they would be difficult to replace.”
- “Licensing related risks” involve dependence on another company’s technology licensing or a government license to operate the business. For example: “If we are unable to renew our licenses or otherwise lose our licensed rights, we may have to stop selling products or we may lose competitive advantage.”

- "Suppliers risks" are any risks related to upstream suppliers, including OEM manufacturers. For example: "A change in sales strategy by the company's suppliers could adversely affect the company's sales or earnings."
- "Input prices risks" are any risks that the input prices (raw material prices) may go up. For example: "Our inability to pass through increases in costs and expenses for raw materials and energy, on a timely basis or at all, could have a material adverse effect on the margins of our products."
- "Rely on few large customers" refers to a high concentration on few large customers. For example: "Our sales could be negatively impacted if one or more of our key customers substantially reduce orders for our products."
- "Competition risks" indicate that the industry is competitive, and the company faces strong or increasing competition. For example: "We compete in distribution industries that are highly competitive and we may not be able to compete successfully."
- "Industry is cyclical" means the industry experiences periodic fluctuations. For example: "We operate in an industry that is cyclical and that has periodically experienced significant year-to-year fluctuations in demand for vehicles."
- "Volatile demand and results" refer to situations where demand and/or financial results are volatile and unpredictable. For example: "Our future revenue, gross margins, operating results and net income are difficult to predict and may materially fluctuate."
- "Volatile stock price risks" indicate that the target company's stock price is volatile. For example: "The price of our common stock has fluctuated widely in the past and may fluctuate widely in the future."
- "New product introduction risks" involve potential delays or failures in new product introduction, or the reliance on new product introduction for the company's success. For example: "Our success depends on our ability to successfully develop and commercialize additional pharmaceutical products."
- "Downstream risks" are risks associated with distributors or retailers. For example: "We face a number of risks related to our product sales through intermediaries."
- "Credit risks" are related to the ability to meet debt obligations, maintain credit agreements, or preserve credit ratings. This includes the danger of breaching loan covenants, losing access to credit facilities, or facing increased costs of borrowing. For example: "If we fail to comply with the covenants in our credit agreements, our

debt could become due immediately, which would have a material adverse effect on our financial condition."

- "Accounting risks" are associated with financial reporting and internal controls, such as the failure to maintain effective internal controls over financial reporting, which could lead to inaccurate financial statements or restatements. For example: "If our internal control over financial reporting is not effective, it could lead to a material misstatement of our financial statements that would not be prevented or detected on a timely basis."

- "Tax risks" arise from changes in tax laws, disputes with tax authorities, or the inability to maintain a particular tax status, which could result in increased tax liabilities or penalties. For example: "If the Internal Revenue Service were to challenge our status as a REIT, we could face significant tax liabilities and our ability to make distributions to stockholders could be adversely affected."

- "Investment risk" involves risks inherent in a company's investment portfolio, such as declines in the value of real estate holdings, loan defaults, or poor performance of market-based investments. For example: "A significant decline in the real estate market could adversely affect the value of our investment portfolio and our financial results."

- "Cost risks" are risks that operational, service, or regulatory costs will increase unexpectedly and cannot be passed on to customers, thereby reducing profit margins. For example: "Rising costs for services and increased industry regulation may compress our profit margins if we are unable to offset them with higher prices or improved efficiency."

- "Cybersecurity risks" are risks of unauthorized access, cyber attacks, or data breaches that could lead to operational disruption, financial loss, theft of intellectual property, or reputational damage. For example: "A significant cybersecurity breach could disrupt our operations, result in the theft of confidential data, and subject us to substantial legal and financial penalties."

- "Insurance and coverage risk" involves risks that an organization is either underinsured, faces gaps in its insurance coverage, or experiences a failure of its insurers, which could leave it financially exposed to losses from operational disruptions, legal liabilities, health care, or property damage. For example: "If our business interruption insurance coverage is insufficient to cover a prolonged shutdown of our primary manufacturing facility, our financial stability and ability to meet customer demand could be severely impacted."

- “Product risk”: The potential for a company’s products or services to have defects, fail to perform as intended, or cause harm, leading to liability claims, reputational damage, recalls, and financial loss. This also includes the risk that a product becomes obsolete or fails to meet market needs.
- “Risk of laggards”: The risk that the company will fall behind competitors or fail to adapt to technological changes, market shifts, or new industry standards, leading to a loss of competitive advantage, market share, or relevance.
- “Not a risk type” refers to text that is not discussing risks or risks that cannot be labeled.

G Additional Empirical Results

G.1 Section 4

Table G.14 shows the Fama-MacBeth cross-section results that exclude companies from tech-related sectors.
Figure G.10 shows the cumulative returns of AI premiums, estimated using companies from non-tech-related sectors.
Table G.15 presents the cross-section results between 2012-2017.

G.2 Section 5

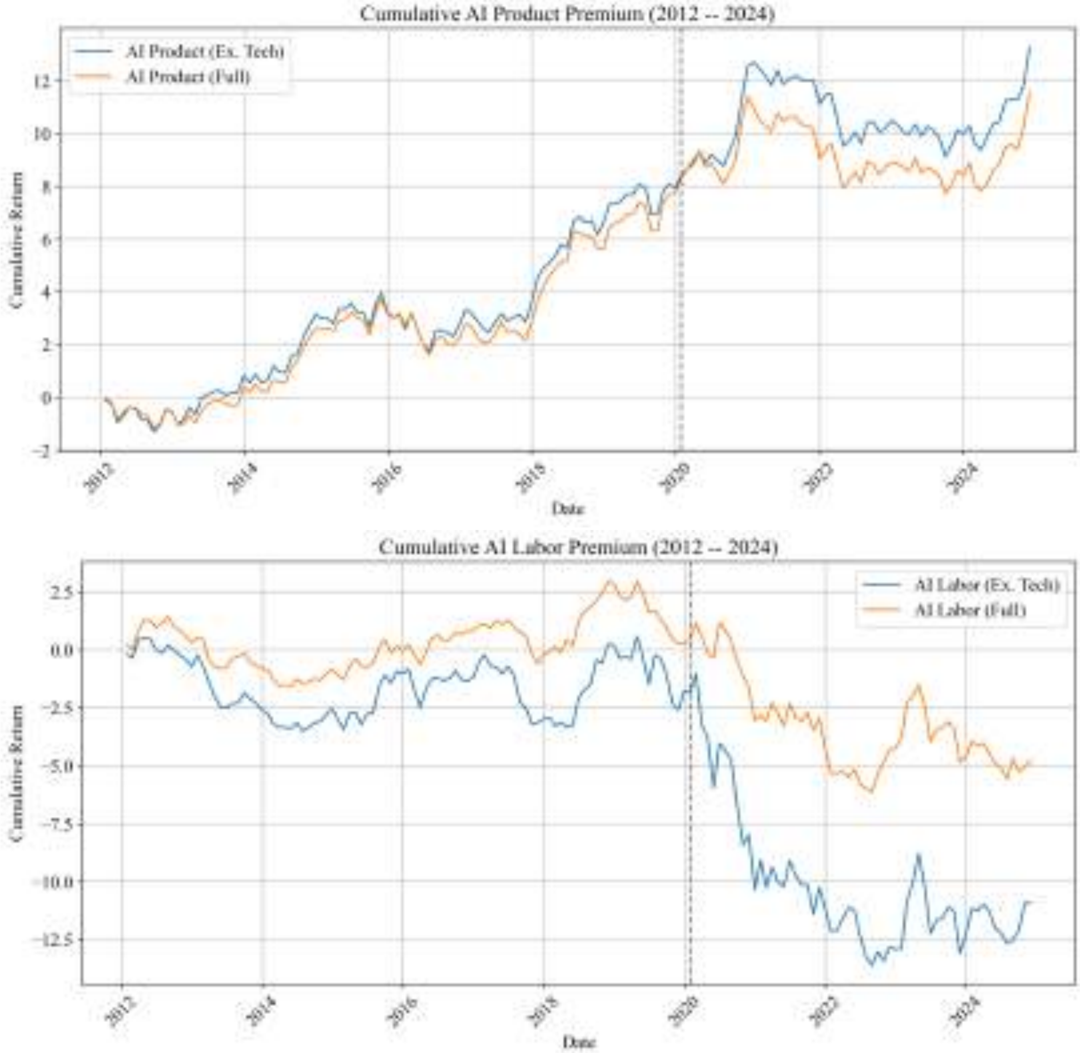
Table G.16 shows the topics and their relative importance over time for companies with high AI risk exposure and high AI product exposure.
Table G.17 shows the topics and their relative importance over time for companies with high AI risk exposure and low AI product exposure.
Table G.18 shows excerpts of Risk Factor section from companies with high AI risk exposure.
Table G.19 shows the results with additional variables and their interaction with the AI exposures.

G.3 Section 6

Table G.20 provides the annual count of firms and average exposure levels for eight distinct AI technology categories from 2011 to 2023.
Table G.21 shows the portfolio sorting results using AI exposure measures of different technologies.
Table G.22 shows the results of Fama-MacBeth regressions using five different alternative measures of AI exposure.

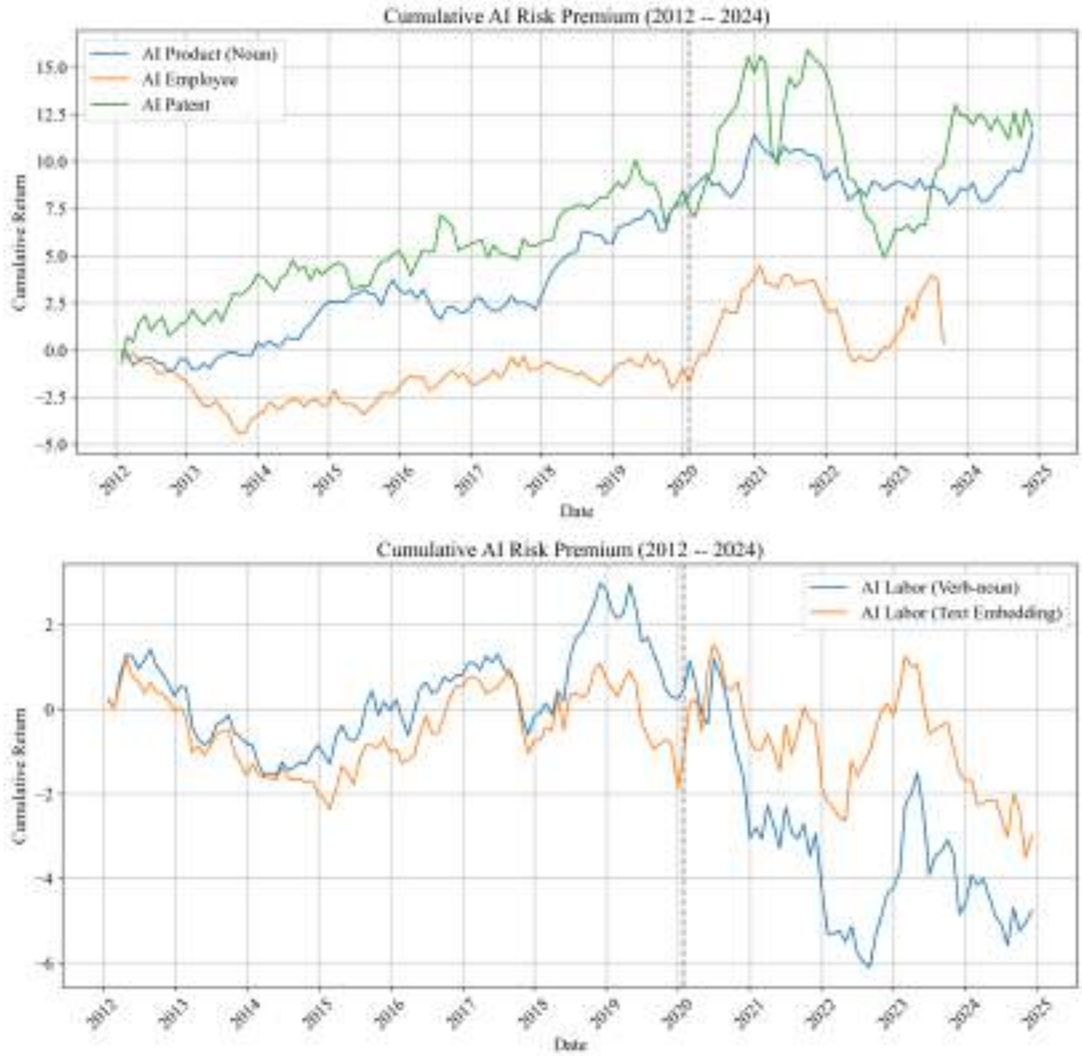
Figure G.11 shows the cumulative value of AI premium of different AI exposure measures.

Figure G.10: Cumulative Returns of AI Premium (Ex. Tech Sector)



Note: This figure shows the cumulative value of the AI premium under different AI measures. We exclude companies in the technology-related sector, with NAICS two-digit codes equal to 51 and 54, as well as the "Magnificent Seven" companies: Apple (AAPL), Amazon (AMZN), Alphabet (GOOG), Meta Platforms (META), Microsoft (MSFT), Nvidia (NVDA), and Tesla (TSLA). The upper graph compares the AI product premium with and without technology-related companies, while the lower graph shows the comparison for the AI labor premium.

Figure G.11: Cumulative Returns AI Premiums from Alternative AI Exposure Measures



Note: This figure shows the cumulative value of AI premiums from different AI exposures. The upper panel shows the results with different AI product exposure. The lower panel shows the results with AI labor exposures. The sample ranges from January 2012 to December 2024. For the AI employee measure, we have sample ranging until mid-2023.

Table G.8: Variable Definitions

Name	Unit	Description
Market Value (MV)	\$1,000,000	The market value of a security expressed in millions, calculated by multiplying the number of shares and security price.
Book-to-Market Value (BM)	-	The balance sheet value of the ordinary (common) equity in the company divided by the total market value.
Operation Profit (oiadp)	\$1,000,000	Operating income represents the difference between sales and total operating expenses.
Total Asset (at)	\$1,000,000	Total assets represent the sum of total current assets, long term receivables, investment in unconsolidated subsidiaries, other investments, net property plant and equipment and other assets.
Operation Profit to Total Asset (OP)	%	Operation profit as a percentage of total assets
Asset Growth (INV)	-	Firm's asset growth, defined as the ratio between total asset of year $m-1$ and year $m-2$.
Leverage (LEV)	%	(Short Term Debt & Current Portion of Long Term Debt + Long Term Debt) / Total Assets * 100.
Firm Volatility (VOL)	%	Historical daily volatility over the past 1 year (250 tradings days).
Property, Plant and Equipment (ppegt)	\$1,000,000	Gross property, plant and equipment less accumulated reserves for depreciation, depletion and amortization.
Cash holdings (CASH)	%	Cash and short-term investment (che) as a percentage of total assets (at).
Illiquidity (ILLIQ)	-	The ratio of the daily absolute stock return to the daily trading volume averaged within the month *1,000,000.
Tangibility (TAN)	%	Property, plant and equipment divided as a percentage of total assets.
Research & Development Expense (RD)	%	R&D expense as a percentage of total sales, imputed with zero if R&D expense data are not available.

Note: The table reports the definition and description of variables used in the Fama-MacBeth regressions.

Table G.9: Top 25 Nouns with Highest Frequency from Patent Universe

Year	2009		2015		2023	
Rank	Noun	Percentage	Noun	Percentage	Noun	Percentage
1	time	22.14%	information	22.33%	information	21.53%
2	control	21.16%	image	19.67%	plurality	19.98%
3	input	20.67%	device	18.52%	user	19.09%
4	set	19.22%	plurality	17.42%	image	18.65%
5	memory	18.52%	user	16.39%	methods	14.86%
6	process	18.45%	apparatus	13.66%	systems	12.53%
7	number	18.07%	invention	13.08%	apparatus	11.92%
8	output	17.80%	computer	12.01%	computer	11.17%
9	network	17.74%	methods	10.63%	set	11.12%
10	value	17.68%	set	10.10%	network	10.55%
11	signal	17.67%	time	9.70%	response	10.41%
12	unit	16.21%	systems	9.53%	time	9.62%
13	embodiment	16.05%	input	9.29%	input	9.61%
14	program	15.65%	network	8.95%	processor	9.07%
15	object	15.13%	embodiment	8.90%	model	8.42%
16	use	14.51%	value	8.50%	example	8.36%
17	response	14.35%	unit	8.49%	unit	8.16%
18	operation	13.74%	control	7.95%	control	8.15%
19	application	13.44%	response	7.92%	application	8.15%
20	interface	12.75%	program	7.71%	value	7.74%
21	processor	12.30%	application	7.70%	embodiments	7.70%
22	function	12.02%	images	7.65%	devices	7.48%
23	order	11.95%	process	7.27%	images	7.35%
24	access	11.91%	number	7.24%	memory	7.34%
25	display	11.76%	memory	6.93%	interface	7.31%

Note: The table shows the top 25 nouns pairs with the highest frequency percentage (the number of patent that contains the noun pair divided by total number of patents). We consider 2009, 2015 and 2023. The patent universe for each year are patent of past ten years (including the current year).

Table G.10: Examples of Verb-noun Pairs from Texts

Example	Text	Verb-Noun Pairs
Patent 1	Disclosed herein are embodiments of systems, methods, and products comprises a server, which provides trigger-based personalized sales outreach based on a user's request. The request comprises a list of target companies/people. The server scans a variety of sources by web crawling the sources' web documents and finds news items relevant to the target companies/people. The server determines an importance score for each news item that measures the probability of a sales representative being interested in the news item. The server applies a set of filters comprising a network of neural networks to filter out false positives. The analytic server determines a relevancy score for each news item. The server generates a GUI to display the news items satisfying certain thresholds. Based on the user's selection on the news items, the server uses the news items as triggers of outreach and generates an electronic message template prepopulated with the news items.	[('provide', 'outreach'), ('crawl', 'document'), ('find', 'item'), ('determine', 'score'), ('apply', 'set'), ('filter', 'positive'), ('determine', 'score'), ('generate', 'gui'), ('display', 'item'), ('satisfy', 'threshold'), ('use', 'item'), ('generate', 'template')]
Patent 2	A system and method for capturing value preference based data from a requestor of one or more digital content presentations during an access sequence and transforming the captured data into a useful tool enabling the content provider to modify the user experience prior to or subsequent to granting access to the requested content resulting in a more efficient and tailored interaction between the requestor and the content provider.	[('capture', 'preference'), ('transform', 'data'), ('enable', 'provider'), ('modify', 'experience'), ('grant', 'access')]
Patent 3	A computer implemented method including receiving, by a monitoring system that is configured to monitor a property and from an electronic pool device that is configured to monitor a swimming pool at the property, sensor data, analyzing, by the monitoring system, the sensor data, based on analyzing the sensor data, generating, by the monitoring system, an instruction to activate a camera of the electronic pool device, providing, by the monitoring system to the electronic pool device, the instruction to activate the camera, receiving, by the monitoring system from the electronic pool device, image data, analyzing, by the monitoring system, the image data, based on analyzing the image data, identifying a monitoring system action to perform, and performing the monitoring system action.	[('receive', 'data'), ('monitor', 'property'), ('monitor', 'pool'), ('analyze', 'data'), ('activate', 'camera'), ('activate', 'camera'), ('receive', 'data'), ('analyze', 'data'), ('identify', 'action'), ('perform', 'action')]
Task 1	Develop sustainability reports, presentations, or proposals for supplier, employee, academia, media, government, public interest, or other groups.	[('develop', 'report')]
Task 2	Write and distribute financial or environmental impact reports.	[('write', 'report'), ('distribute', 'report')]
Task 3	Direct and coordinate organization's financial and budget activities to fund operations, maximize investments, and increase efficiency.	[('fund', 'operation'), ('maximize', 'investment'), ('increase', 'efficiency'), ('direct', 'activity'), ('coordinate', 'activity')]

Note: The table shows the examples of verb-noun pairs extracted from given text. We provide three examples of patent abstract and task description.

Table G.11: Top 25 Verb-noun Pairs with Highest Frequency from Patent Universe

2009		2015		2023	
Verb-noun Pair	Perc.	Verb-noun Pair	Perc.	Verb-noun Pair	Perc.
('receive', 'data')	1.85%	('receive', 'data')	2.31%	('receive', 'data')	3.28%
('receive', 'request')	1.74%	('receive', 'request')	2.28%	('receive', 'request')	2.83%
('provide', 'method')	1.71%	('generate', 'image')	1.57%	('perform', 'operation')	1.80%
('provide', 'system')	1.50%	('provide', 'method')	1.55%	('generate', 'image')	1.79%
('generate', 'image')	1.33%	('receive', 'information')	1.49%	('capture', 'image')	1.68%
('use', 'information')	1.32%	('capture', 'image')	1.47%	('receive', 'information')	1.66%
('provide', 'information')	1.30%	('use', 'information')	1.36%	('provide', 'method')	1.63%
('store', 'data')	1.26%	('provide', 'system')	1.35%	('use', 'data')	1.62%
('receive', 'information')	1.24%	('perform', 'operation')	1.33%	('generate', 'data')	1.61%
('use', 'data')	1.22%	('use', 'data')	1.30%	('receive', 'input')	1.39%
('capture', 'image')	1.18%	('generate', 'data')	1.27%	('use', 'model')	1.39%
('perform', 'operation')	1.17%	('provide', 'information')	1.26%	('use', 'information')	1.27%
('generate', 'data')	1.15%	('store', 'data')	1.22%	('provide', 'system')	1.24%
('generate', 'signal')	1.10%	('receive', 'signal')	1.14%	('store', 'data')	1.21%
('receive', 'signal')	1.08%	('receive', 'input')	1.06%	('receive', 'signal')	1.21%
('store', 'information')	0.88%	('generate', 'signal')	1.04%	('obtain', 'data')	1.14%
('process', 'data')	0.87%	('display', 'image')	0.89%	('provide', 'information')	1.14%
('display', 'image')	0.81%	('process', 'data')	0.86%	('obtain', 'information')	1.03%
('receive', 'input')	0.80%	('obtain', 'information')	0.86%	('generate', 'signal')	1.01%
('enable', 'user')	0.79%	('store', 'information')	0.81%	('process', 'data')	0.99%
('determine', 'value')	0.77%	('determine', 'value')	0.80%	('obtain', 'image')	0.94%
('obtain', 'information')	0.76%	('use', 'device')	0.79%	('receive', 'image')	0.94%
('calculate', 'value')	0.73%	('enable', 'user')	0.79%	('determine', 'value')	0.93%
('use', 'model')	0.69%	('obtain', 'image')	0.74%	('display', 'image')	0.89%
('use', 'system')	0.67%	('obtain', 'data')	0.73%	('use', 'device')	0.89%

Note: The table shows the top 25 verb-noun pairs with the highest frequency percentage (the number of patent that contains the verb-noun pair divided by total number of patents). We consider 2009, 2015 and 2023. The patent universe for each year are patent of past ten years (including the current year).

Table G.12: Tasks with the Highest Exposures

Task ID	Task Description
11403	Compare data with source documents, or re-enter data in verification format to detect errors.
3665	Observe workers operating equipment or performing tasks to determine time involved and fatigue rate using timing devices.
1192	Perform searches for qualified candidates according to relevant job criteria, using computer databases, networking, Internet recruiting resources, cold calls, media, recruiting firms, and employee referrals.
15905	Analyze logistics data, using methods such as data mining, data modeling, or cost or benefit analysis.
3487	Assist users to diagnose and solve data communication problems.
16095	Enter data into Geographic Information Systems (GIS) databases, using techniques such as coordinate geometry, keyboard entry of tabular data, manual digitizing of maps, scanning or automatic conversion to vectors, or conversion of other sources of digital data.
16265	Analyze clinical or survey data, using statistical approaches such as longitudinal analysis, mixed-effect modeling, logistic regression analyses, and model-building techniques.
16283	Analyze clinical data using appropriate statistical tools.
15298	Process data for analysis, using computers.
5392	Analyze available data and consult with other scientists to determine parameters of experimentation and suitability of analytical models.
16613	Produce images of objects on film, using radiographic techniques.
16716	Perform diagnostic analyses of processing steps, using analytical or metrological tools, such as microscopy, profilometry, or ellipsometry devices.
174	Program and use computers to store, process, and analyze data.
10903	Perform complex calculations as part of the analysis and evaluation of data, using computers.
16872	Collect and analyze data to determine environmental conditions and restoration needs.
16868	Create environmental models or simulations, using geographic information system (GIS) data and knowledge of particular ecosystems or ecological regions.
18270	Manage or analyze data obtained from remote sensing systems to obtain meaningful results.
5437	Seek and provide information to help companies determine their position in the marketplace.
16999	Collect geospatial data, using technologies such as aerial photography, light and radio wave detection systems, digital satellites, or thermal energy systems.
9188	Collect information about clients, using techniques such as testing, interviewing, discussion, or observation.

Note: The table shows the top 20 tasks with the highest exposure (cosine similarity) over the sample period.

Table G.13: Occupations with Extreme on AI Labor Exposures

Period	Top 5 Occupations	Bottom 5 Occupations
2009-2014	Market Research Analysts and Marketing Specialists	Helpers--Painters, Paperhangers, Plasterers, and Stucco Masons
	Credit Authorizers, Checkers, and Clerks	Musicians and Singers
	Receptionists and Information Clerks	Tapers
	Nuclear Engineers	Maids and Housekeeping Cleaners
	Telephone Operators	Butchers and Meat Cutters
2015-2019	Market Research Analysts and Marketing Specialists	Tapers
	Credit Authorizers, Checkers, and Clerks	Musicians and Singers
	Telephone Operators	Helpers--Painters, Paperhangers, Plasterers, and Stucco Masons
	Nuclear Engineers	Maids and Housekeeping Cleaners
	Telephone Operators	Helpers--Carpenters
2020-2023	Credit Authorizers, Checkers, and Clerks	Tapers
	Nuclear Engineers	Musicians and Singers
	Private Detectives and Investigators	Floor Sanders and Finishers
	Chemical Engineers	Helpers--Painters, Paperhangers, Plasterers, and Stucco Masons
	Forest and conservation technicians	Maids and Housekeeping Cleaners

Note: This table shows the occupations with the highest and lowest labor exposure of different periods. For the results of 2009-2014, we have extended the construction of AI labor measure back to 2009.

Table G.14: Fama-MacBeth Regressions (Ex. Tech)

Dep.	R					
Periods	2012-2024			2012-2019		
Models	(1)	(2)	(3)	(4)	(5)	(6)
$AI_{Product}$	0.0841**		0.0856**	0.0802**		0.0825**
	(2.07)		(2.17)	(2.03)		(2.11)
AI_{Labor}		-0.0664	-0.0704		-0.0185	-0.0183
		(-1.14)	(-1.24)		(-0.37)	(-0.37)
Observations	335037	359136	332637	212961	224341	211469
R^2	0.10	0.10	0.10	0.09	0.09	0.09
Controls	Y	Y	Y	Y	Y	Y
Sector Dummies	Y	Y	Y	Y	Y	Y

Note: This table reports Fama-MacBeth regressions of individual stock returns on AI exposures and other firm characteristics. We exclude companies in the tech-related sector, with NAICS two-digit codes equal to 51 and 54. We also exclude “magnificent seven” stocks: Apple (AAPL), Amazon (AMZN), Alphabet (GOOG), Meta Platforms (META), Microsoft (MSFT), Nvidia (NVDA), and Tesla (TSLA) * Statistical significance at the 10% level;** Statistical significance at the 5% level;*** Statistical significance at the 1% level.

Table G.15: Fama-MacBeth Regressions (2012-2017)

Dep.		R	
Models	(1)	(2)	(3)
$AI_{Product}$	0.0374		0.0394
	(0.98)		(1.03)
AI_{Labor}		0.0052	-0.0025
		(0.16)	(-0.07)
Observations	185911	197214	184798
R^2	0.08	0.08	0.08
Controls	Y	Y	Y
Sector Dummies	Y	Y	Y

Note: This table reports Fama-MacBeth regressions of individual stock returns on the AI adoption measure and other firm characteristics. The sample period ranges from January 2012 to December 2017. * Statistical significance at the 10% level;** Statistical significance at the 5% level;*** Statistical significance at the 1% level.

Table G.16: Topics for High-AI Risk and High-AI Product (Group 1)

Year/Rank	2016	2017	2018	2019	2020	2021	2022	2023
1	Financial condition 8.99%	Financial condition 7.76%	Financial condition 7.91%	**New product introduction** 16.29%	**New product introduction** 7.15%	Financial condition 7.78%	**New product introduction** 10.08%	Macroeconomic 8.44%
2	Competition 6.45%	**New product introduction** 7.20%	Infrastructure 5.61%	Financial condition 7.60%	Financial condition 5.90%	**New product introduction** 6.28%	Financial condition 7.90%	Regulation changes 5.74%
3	**New product introduction** 5.61%	Regulation changes 5.19%	Competition 5.04%	Regulation changes 5.49%	Infrastructure 4.67%	Funding 4.98%	Funding 5.72%	Merger & Acquisition 5.41%
4	Regulation changes 4.91%	Human resource 5.19%	Intellectual property 4.99%	Competition 5.17%	Funding 4.65%	Competition 4.84%	Competition 4.45%	Intellectual property 4.61%
5	Funding 4.85%	Macroeconomic 4.65%	Funding 4.53%	Cost 4.04%	Competition 3.95%	Intellectual property 4.23%	Regulation changes 4.20%	Competition 4.51%
6	Intellectual property 4.54%	Funding 4.44%	Macroeconomic 4.45%	Infrastructure 3.98%	Intellectual property 3.95%	Cost 4.09%	Intellectual property 3.67%	Human resource 3.81%
7	**Cybersecurity risk** 4.05%	Licensing related 3.90%	Rely on few large customers 4.32%	International 3.31%	Regulation changes 3.91%	Macroeconomic 3.85%	**Cybersecurity risk** 3.22%	Volatile demand and results 3.52%
8	Merger & Acquisition 4.01%	Suppliers 3.34%	Potential defects in products 4.05%	Funding 3.30%	Human resource 3.34%	Infrastructure 3.69%	Licensing related 3.20%	International 3.39%
9	Potential defects in products 3.39%	Intellectual property 3.19%	Regulation changes 3.70%	Volatile stock price 3.19%	**Cybersecurity risk** 3.26%	Suppliers 3.61%	Suppliers 3.16%	**Cybersecurity risk** 3.26%
10	Disruption of operations 3.31%	Volatile stock price 2.89%	Licensing related 3.52%	Licensing related 3.09%	Suppliers 3.05%	Regulation changes 3.57%	Volatile stock price 3.05%	Disruption of operations 3.06%

Table G.17: Topics for High-AI Risk and Low-AI Product (Group 2)

Year/Rank	2016	2017	2018	2019	2020	2021	2022	2023
1	Financial condition 8.71%	Competition 9.86%	Financial condition 10.77%	Intellectual property 11.03%	Competition 12.44%	Financial condition 8.42%	Financial condition 10.46%	Financial condition 10.25%
2	Macroeconomic 6.60%	Disruption of operations 7.94%	Competition 10.05%	Volatile demand and results 9.23%	Financial condition 10.34%	Competition 7.07%	Regulation changes 5.64%	Suppliers 6.86%
3	Regulation changes 5.80%	Regulation changes 7.47%	Regulation changes 7.59%	Financial condition 6.28%	Regulation changes 7.08%	Macroeconomic 4.53%	**Cybersecurity risk** 4.56%	Regulation changes 6.10%
4	Competition 5.18%	Downstream 5.22%	Cost 5.41%	Regulation changes 4.29%	Risk of laggards 4.39%	Funding 3.92%	Cost 4.00%	Macroeconomic 5.63%
5	Cost 4.69%	Funding 4.22%	**Cybersecurity risk** 4.37%	Accounting 3.92%	Infrastructure 3.78%	Regulation changes 3.85%	Product risk 3.97%	Competition 5.49%
6	**Cybersecurity risk** 4.54%	Merger & Acquisition 4.18%	Macroeconomic 3.75%	Infrastructure 3.67%	Volatile stock price 3.77%	Merger & Acquisition 3.38%	Suppliers 3.74%	Cost 4.30%
7	Funding 4.00%	Macroeconomic 3.78%	Merger & Acquisition 3.65%	Merger & Acquisition 3.47%	**New product introduction** 3.48%	Human resource 3.22%	Competition 3.56%	Human resource 3.52%
8	Risk management 3.79%	Financial condition 3.48%	**New product introduction** 3.50%	Human resource 3.47%	Merger & Acquisition 3.38%	Accounting 3.21%	Accounting 3.30%	Potential/Ongoing Lawsuits 3.47%
9	Intellectual property 3.48%	Infrastructure 3.43%	Funding 3.32%	Risk of laggards 3.43%	Licensing related 3.15%	Volatile stock price 3.04%	Funding 3.27%	Merger & Acquisition 3.32%
10	Disruption of operations 3.17%	International 3.42%	Infrastructure 3.10%	Input prices 3.03%	Funding 3.02%	**Cybersecurity risk** 2.94%	Credit 3.20%	Funding 3.31%

Table G.18: Excerpts of Risk Factor Texts related to Technology and AI

Name	Year	Product	AI Risk	NAICS	Excerpts
Farmers Capital Bank Corporation	2016	0.000	0.220	522110	-A failure in or breach, including cyber attacks, of the Company's operational or security systems, ...increase its costs and cause losses. -Technology and other changes now allow many consumers to complete financial transactions without using banks.
Key Energy Services, Inc.	2020	0.000	0.221	2111	-We may not be successful in implementing and maintaining technology development and enhancements. -New technology may cause us to become less competitive. -Conservation measures and technological advances could reduce demand for oil and natural gas. -Our operations may be subject to **cyber-attacks** ....
Independent Bank Group, Inc.	2022	0.000	0.174	522110	The Company's accounting estimates and risk management programs rely on analytical and forecasting models. The Company may experience system failure or cybersecurity breaches. -The success of the Company is dependent upon the strength of its recruitment efforts, as well as its succession plans and procedures. –The Company must effectively manage the need for technological change. -The Company may be subject to data processing failures, control failures and fraud.
Synchronoss Technologies, Inc.	2023	0.328	0.309	518210	-The markets in which we market and sell our products and services are highly competitive, and if we do not adapt to rapid technological change, we could lose customers or market share. -Technological development by others may impact the competitiveness of our products in the marketplace.
Rigetti Computing, Inc.	2023	0.298	0.172	334413	-Although we currently believe that quantum machine learning for finance is poised to be an early domain of quantum ..., the risks associated with developing a product that can compute algorithms that scale efficiently to real-world size applications -We may not be able to reduce the cost of developing our quantum computers
Recursion Pharmaceuticals	2023	0.312	0.274	325414	-Our platform depends upon the continuous, effective, and reliable operation of our software, hardware, databases, and related tools and functions, as well as the integrity of our data. -Our proprietary software tools, hardware, and data sets are inherently complex. -We have from time to time found defects, vulnerabilities, or other errors in our software and hardware that produce... new errors with our software and hardware may be detected in the future. -Issues relating to the use of new and evolving technologies such as AI and machine learning may cause us to experience brand or reputational harm... -Known risks of AI currently include inaccuracy, bias, toxicity, intellectual property infringement or misappropriation, data privacy and cybersecurity issues, and data provenance disputes. -Regulatory and legislative developments related to the use of AI could adversely affect our use of such technologies in our products, services, and business...

Table G.19: Fama-MacBeth Regressions with Additional Variables

Dep.	R							
Periods	2012-2020							
Models	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$AI_{Product}$	0.1112***	0.1199***	0.1060**	0.1050**	0.1056**	0.1052**	0.1063**	0.1040**
	(2.88)	(3.09)	(2.46)	(2.42)	(2.50)	(2.45)	(2.50)	(2.42)
$Digital$	-0.0239	-0.0307						
	(-0.44)	(-0.57)						
Political Risk			0.0151	0.0156				
			(0.66)	(0.68)				
Non-Political Risk					-0.0386	-0.0422		
					(-1.55)	(-1.58)		
Overall Risk							-0.0185	-0.0167
							(-0.48)	(-0.42)
$Digital \times AI_{Product}$		0.0412						
		(1.09)						
Political Risk $\times$ $AI_{Product}$				0.0167				
				(0.90)				
Non-Political Risk$\times AI_{Product}$						-0.0019		
						(-0.11)		
Overall Risk$\times AI_{Product}$								-0.0052
								(-0.21)
Observations	276022	276022	217635	217635	217635	217635	217635	217635
R^2	0.08	0.08	0.11	0.11	0.11	0.11	0.11	0.11
Controls	Y	Y	Y	Y	Y	Y	Y	Y
Sector Dummies	Y	Y	Y	Y	Y	Y	Y	Y

Note: This table reports Fama-MacBeth regressions of individual stock returns on AI exposures and additional variables.
* Statistical significance at the 10% level;** Statistical significance at the 5% level;*** Statistical significance at the 1% level.

Table G.20: Summary Statistics for AI Technology Exposure

Panel A: Number of companies with non-zero AI technology exposure

Year	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Machine Learning	164	182	190	200	188	199	191	202	242	260	302	319	334
Evolutionary Computation	139	159	157	152	144	157	157	154	186	195	207	215	226
Natural Language Processing	175	187	199	211	214	216	209	207	220	243	236	263	253
Speech	111	132	135	141	140	154	141	141	143	153	167	167	179
Computer Vision	242	271	272	288	291	288	285	290	318	329	317	334	347
Planning and Control	382	402	418	438	449	443	420	439	465	466	497	502	497
Knowledge Processing	241	257	261	289	286	280	279	306	328	322	335	340	354
AI Hardware	303	320	312	341	331	329	328	334	379	358	379	400	412

Panel B: Average AI technology exposure for each technology in each year (%)

Year	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Machine Learning	1.20	1.71	1.55	1.73	1.50	2.01	1.30	1.89	2.38	2.99	4.09	4.90	5.05
Evolutionary Computation	0.60	1.04	0.57	0.53	0.48	0.58	0.50	0.78	0.98	1.00	1.31	1.10	1.38
Natural Language Processing	2.45	2.48	2.79	2.65	2.62	3.29	2.74	2.86	3.37	3.30	3.54	3.89	3.76
Speech	1.04	1.15	1.21	0.96	1.17	1.05	0.97	0.87	1.08	1.05	1.28	1.11	1.11
Computer Vision	2.60	2.68	2.80	2.97	3.12	3.16	3.31	3.65	4.05	4.23	4.32	5.08	4.96
Planning and Control	7.12	8.06	8.37	7.97	7.65	8.20	8.07	9.29	9.94	10.46	12.30	12.39	11.91
Knowledge Processing	3.60	3.82	3.39	4.30	4.13	4.55	4.27	5.24	5.42	5.97	5.82	6.56	6.78
AI Hardware	3.53	3.86	3.81	4.15	3.79	4.89	3.98	4.60	6.06	5.96	7.01	7.86	8.35

Note: This table shows the number of firms with non-zero AI technology exposure (Panel A) and the average AI technology exposure in percentage (Panel B) across eight AI technology categories from 2011 to 2023 (fiscal year).

Table G.21: Portfolio Sorts with Different AI Technologies

		Mean	CAPM	FF3	FF3+UMD	FF5	FF5+UMD
Machine Learning	H-L(VW)	0.479**	0.373	0.209	0.256	0.277	0.308
		(2.16)	(1.63)	(1.02)	(1.24)	(1.40)	(1.54)
	H-L(EW)	0.579***	0.549***	0.367**	0.369**	0.397***	0.387***
		(3.40)	(3.10)	(2.52)	(2.49)	(2.76)	(2.65)
Evolutionary Computation	H-L(VW)	0.4256*	0.2825	0.1439	0.2133	0.2216	0.2691
		(2.01)	(1.30)	(0.75)	(1.12)	(1.25)	(1.51)
	H-L(EW)	0.5236***	0.5252**	0.3070*	0.3904**	0.3281*	0.4012**
		(2.62)	(2.52)	(1.81)	(2.34)	(1.94)	(2.40)
Natural Language Processing	H-L(VW)	0.6898***	0.6036***	0.4701**	0.5166**	0.5529***	0.5825***
		(3.19)	(2.69)	(2.32)	(2.53)	(2.88)	(3.00)
	H-L(EW)	0.5305***	0.4934***	0.3288**	0.3541**	0.3637**	0.3844**
		(2.94)	(2.62)	(2.00)	(2.13)	(2.22)	(2.31)
Speech	H-L(VW)	0.5774***	0.4498**	0.3031	0.3614*	0.3540**	0.3931**
		(2.80)	(2.12)	(1.63)	(1.93)	(1.98)	(2.17)
	H-L(EW)	0.4743***	0.4611***	0.2879**	0.3228**	0.2843**	0.3085**
		(2.95)	(2.75)	(2.09)	(2.32)	(2.08)	(2.23)
Computer Vision	H-L(VW)	0.8433***	0.7476***	0.5418**	0.5673**	0.5966***	0.5973***
		(3.17)	(2.71)	(2.31)	(2.39)	(2.63)	(2.59)
	H-L(EW)	0.2394	0.1953	0.0873	0.1083	0.0951	0.1137
		(1.74)	(1.36)	(0.66)	(0.81)	(0.71)	(0.84)
Planning and Control	H-L(VW)	0.3716*	0.3281	0.2558	0.2402	0.3531*	0.3298*
		(1.94)	(1.64)	(1.29)	(1.19)	(1.89)	(1.74)
	H-L(EW)	0.5852***	0.5663***	0.4035**	0.3769**	0.4633**	0.4316**
		(2.93)	(2.72)	(2.13)	(1.96)	(2.49)	(2.29)
Knowledge Processing	H-L(VW)	0.5726***	0.5259**	0.4052**	0.4163**	0.4909**	0.4894**
		(2.75)	(2.43)	(1.99)	(2.02)	(2.53)	(2.49)
	H-L(EW)	0.5642***	0.5058**	0.3516*	0.3438*	0.4149**	0.3992**
		(2.89)	(2.49)	(1.91)	(1.83)	(2.30)	(2.18)
AI Hardware	H-L(VW)	0.6696**	0.5273*	0.3488	0.3818	0.4567**	0.4667**
		(2.51)	(1.92)	(1.41)	(1.53)	(1.97)	(1.99)
	H-L(EW)	0.6510***	0.5770***	0.4371**	0.4320**	0.4803**	0.4698**
		(3.24)	(2.77)	(2.28)	(2.22)	(2.51)	(2.42)

Note: The sample period ranges from January 2012 to December 2024.
* Statistical significance at the 10% level;** Statistical significance at the 5% level;*** Statistical significance at the 1% level.

Table G.22: Fama-MacBeth Regressions with Alternative AI Exposures

Dep.	R				
Periods	2012-2020				
Models	(1)	(2)	(3)	(4)	(5)
$AI_{Product}$	0.1057**				
	(2.55)				
$AI_{Employee}$		0.0358			
		(0.72)			
AI_{Patent}			0.1362**		
			(2.38)		
AI_{Labor}				-0.0283	
				(-0.59)	
AI^{NLP}_{Labor}					-0.0067
					(-0.21)
Observations	276022	198936	95503	276022	276022
R^2	0.08	0.10	0.10	0.08	0.08
Controls	Y	Y	Y	Y	Y
Sector Dummies	Y	Y	Y	Y	Y

Note: This table reports Fama-MacBeth regressions of individual stock returns on the different AI exposure measures and other firm characteristics. The sample ranges from January 2012 to December 2020.