Data Concerns and Investments in AI/ML: Demand- and Supply-Side Implications

Beatrice (Hyeongyun) Chang *

November 2025

[Preliminary Draft]

Abstract

This paper studies how data privacy and national security concerns shape the internal development and external adoption of AI technologies. On the demand side, we examine whether security-sensitive firms, such as those in healthcare, defense, or government, invest more in internal AI capabilities. Using job postings data, we find that these firms are more likely to hire AI talent in-house. On the supply side, we analyze exit outcomes for AI startups. Startups backed by security-sensitive firms are less likely to exit via IPO or acquisition, suggesting lower external AI adoption. However, cybersecurity-focused startups, closely aligned with these firms' strategic concerns, are more likely to exit, especially when backed by security-sensitive investors. These findings highlight how security concerns shape both internal innovation strategy and the external AI innovation pipeline.

Keywords: *Artificial Intelligence, Machine Learning, Data Privacy, Cybersecurity, Startup Exits, Venture Investments, Innovation Adoption*
JEL Classification: *G24, L86, O31, M15*

*Warrington College of Business, University of Florida. chang.hy@ufl.edu
I thank Joel Houston, Yuehua Tang, and Minmo Gahng for their excellent advice.

1. Introduction

Over the past decade, the adoption of Artificial Intelligence (AI) and Machine Learning (ML) has accelerated at an extraordinary pace. Using the full universe of Lightcast job postings from 2010 to May 2025, we observe a persistent rise in AI-related hiring across U.S. firms, with AI jobs growing as a share of all postings even during periods of macroeconomic slowdown (Figure A.1). Venture investment patterns echo this trend: as shown in Figure A.3, the share of VC dollars allocated to AI has expanded sharply, especially after 2016. These patterns underscore a broad economic shift toward data-driven decision-making, automated workflows, and AI-enabled products.

Yet this rapid diffusion masks a deep heterogeneity in how firms adopt AI technologies. AI systems require access to proprietary or sensitive data, frequent data transfers, and integration with external software infrastructure. For firms operating in healthcare, defense, and regulated segments of financial services, the cost of data exposure—legal, operational, and reputational—is exceptionally high. These "security-sensitive" firms face a fundamentally different optimization problem: whether to acquire external AI tools or to develop AI capabilities internally.

Our first contribution is to show that firms exposed to high data-governance risk disproportionately internalize AI capability building. Using job postings data from Lightcast matched to Compustat financials, we document that security-sensitive firms consistently exhibit higher AI job shares relative to otherwise similar firms. This pattern emerges even after controlling for firm size, profitability, leverage, and Tobin's Q, and persists under year and industry fixed effects. These facts suggest that concerns over confidentiality, compliance, and regulatory scrutiny push certain firms toward "building" rather than "buying" AI, consistent with classic theories of organizational boundaries and incomplete contracting.

However, this internalization on the demand side has profound implications for the supply side of the AI economy. AI/ML service-provider startups depend on access to enterprise clients for adoption and revenue growth. Yet we document striking evidence that B2B-oriented AI startups have experienced worsening performance over time. Figure A.4 shows that the share of B2B AI companies has declined steadily, and Figure A.5 illustrates that their success rates (IPO or high-value M&A relative to total VC invested) have fallen faster than those of other AI startups. We further show that B2B AI startups face lower IPO probabilities and lower exit valuations, and survive less often in Cox hazard models (Figure A.6). These patterns point to a fraying link between enterprise demand and external AI vendors.

Why do B2B AI startups underperform in precisely the period when AI adoption is surging? We argue that their struggle is directly tied to the demand-side internalization documented above. When security-sensitive firms are reluctant to adopt externally built tools—especially tools requiring deep access to proprietary data—the market becomes segmented. Only two types of AI startups tend to succeed: those that resemble incumbent firms closely enough to reduce adoption frictions, and those that specialize in cybersecurity capabilities that directly alleviate data-protection concerns. In Section 6, we show that startups with higher textual similarity to incumbents have higher exit probabilities but lower valuations, while cybersecurity-focused AI startups achieve faster and more valuable exits, especially in regulated markets.

Our study therefore links the demand- and supply-sides of the AI ecosystem through a unified mechanism: data governance risk. On the demand side, sensitive firms internalize AI development. On the supply side, external vendors must either mimic incumbents or specialize in mitigating data risk to overcome adoption barriers. This dual structure helps explain why AI adoption is booming while B2B AI exits are falling, and why the AI startup landscape is increasingly bifurcated.

The remainder of the paper proceeds as follows. Section 2 reviews related literature. Section 3 describes the data and sample construction. Section 4 outlines our empirical framework. Section 5 presents the main results. Section 6 connects demand- and supply-side dynamics. Section 7 discusses robustness. Section 8 considers alternative explanations. Section 9 concludes.

2. Related Literature

Our study contributes to several intersecting strands of literature at the nexus of technological innovation, organizational strategy, and data governance. In particular, we build on research in the following domains: (1) the economics of organizational boundaries and internalization decisions, (2) the measurement and determinants of AI adoption across firms, and (3) the role of cybersecurity, regulation, and compliance in shaping innovation strategies in sensitive sectors.

First, our paper connects to a long-standing theoretical tradition on the boundaries of the firm. The question of whether to "make or buy" has shaped firm strategy for decades, with roots in Coase (1937)'s transaction cost theory and Williamson (1985)'s work on asset specificity and opportunism. These foundational models emphasize that firms internalize certain functions when market transactions become too costly, risky, or inefficient. Building on this, Arora and Gambardella (1990) show that firms are more likely to internalize R&D when knowledge spillovers or appropriation concerns are high, particularly in high-tech environments. Gans and Stern (2003) formalize the idea that downstream commercialization hazards discourage outsourcing. These frameworks suggest that security-sensitive firms may be more likely to internalize AI capability building, particularly those involving sensitive data, as too risky to entrust to third-party vendors. Internal hiring of AI engineers

and data scientists becomes a strategic substitute for outsourcing, offering firms greater control over data handling, IP, and system architecture.

Recent empirical studies have adapted this framework to digital technologies. For example, Casadesus-Masanell and Ricart (2010) argue that technology adoption decisions are deeply embedded in business model design, which reflects both competitive positioning and institutional constraints. In our setting, firms in healthcare, defense, or finance are not just responding to costs, but to legal exposure, reputational risk, and regulatory surveillance. These frictions reshape the cost-benefit calculus of external versus internal AI capability building.

Second, we build on the rapidly growing empirical literature on AI diffusion and workforce transformation. Babina et al. (2024) introduce a novel dataset based on job postings from Lightcast (formerly Burning Glass) to measure firm-level AI investment. They document that AI adopters tend to be larger, more profitable, and more innovative. In related work, Webb (2019) and Brynjolfsson et al. (2023) examine the types of tasks and occupations most affected by AI, showing that AI hiring is highly skewed toward firms with aligned managerial practices and digital infrastructure. These studies emphasize that job postings offer granular, high-frequency insights into how AI capabilities diffuse across the economy. We adopt a similar approach but contribute a novel angle by linking adoption patterns to firms' data governance environments. That is, we do not just examine *whether* firms adopt AI, but *how* institutional and sectoral constraints shape the form of adoption.

Third, we contribute to a growing literature on data regulation, cybersecurity risk, and digital governance. Gal and Rubinfeld (2019) provide an overview of how data protection laws, such as GDPR in the EU or HIPAA in the U.S., constrain firm behavior in data-rich sectors. Acemoglu et al. (2022) and Zuboff (2019) theorize how surveillance capitalism can generate both economic distortions and regulatory frictions, potentially chilling innovation. These studies provide the macro-context in which security-sensitive

firms operate, an environment where legal ambiguity, reputational scrutiny, and data breaches carry enormous costs. For example, a health insurer deploying third-party AI tools risks violating HIPAA; a defense contractor adopting external NLP tools may trigger national security concerns. Our classification of "security-sensitive" firms is designed to capture this elevated exposure and its strategic implications.

Cybersecurity research further reinforces our premise. Cavusoglu et al. (2004) estimate that public firms suffer economically significant valuation losses following data breaches. Gordon et al. (2015) develop an economic model of optimal cybersecurity investment and find that firms systematically underinvest, particularly where enforcement is weak. Romanosky (2016) shows that breach costs vary widely by sector, with the most punitive outcomes in healthcare and finance. Anderson and Moore (2006) highlight the agency problems that arise when data is handled by outside vendors, emphasizing why internalization may be preferred in high-risk settings.

Finally, our research engages with broader theories of digital transformation and heterogeneous firm capability. Bresnahan et al. (2002) argue that IT investments yield returns only when paired with complementary human capital and organizational practices. Our findings, that security-sensitive firms disproportionately hire AI talent, align with this view, suggesting that firms customize their internal workforce to reflect institutional realities. Bloom et al. (2019) emphasize the role of managerial practices in mediating IT-led productivity gains, reinforcing the idea that AI hiring is not just a technical decision but a strategic one.

Together, this body of work provides the conceptual and empirical foundations for our study. By highlighting data sensitivity as a key driver of internal AI investment, we offer a novel explanatory channel for firm-level heterogeneity in the digital era. This framework also motivates our investigation of supply-side dynamics, specifically how data sensitivity shapes exit outcomes for AI startups that rely on enterprise adoption. Taken together, our

analysis connects demand- and supply-side behavior in AI markets, grounded in the dual logics of risk management and strategic complementarity.

3. Data and Hypotheses

3.1. Data Sources and Construction

Our empirical analysis draws on two complementary datasets that capture internal AI capability building by public firms (Hypothesis 1) and external AI adoption through startup exits (Hypotheses 2a and 2b).

Internal AI hiring (Hypothesis 1). We combine U.S. job postings from Lightcast (formerly Burning Glass Technologies) with firm-level financials from Compustat North America. The Lightcast data cover 2010–2025 and include employer names, job titles, ONET occupation codes, NAICS sectors, and detailed job descriptions. Following Babina et al. (2024), we classify a posting as AI-related if it belongs to AI-intensive occupations (e.g., Data Scientists), AI-relevant functional areas (e.g., IT, R&D), or contains keywords such as "deep learning," "neural networks," or "natural language processing." Postings are aggregated to the firm-year level and matched to Compustat using cleaned employer names. Our primary outcome is the AI hiring share, defined as the number of AI-related postings divided by a firm's total postings in a given year. We also examine log AI posting counts and the non-AI hiring share as alternative outcomes.

Startup exits and investor sensitivity (Hypotheses 2a and 2b). To analyze external AI adoption, we use PitchBook data on global VC-backed startups tagged under the Artificial Intelligence & Machine Learning vertical. We retain startups receiving at least one VC deal between 2010 and 2024 and collect investor identities, exit events (IPO and M&A), acquirer characteristics, and funding histories. We classify investors as security-sensitive if they are

affiliated with, backed by, or subsidiaries of firms in healthcare, defense, government, or other data-sensitive industries subject to regulatory oversight (e.g., HIPAA-covered entities). We also flag cybersecurity-focused startups using PitchBook verticals and keyword-based tagging (e.g., "threat detection," "data encryption").

3.2. Hypothesis 1: Demand-Side — Security Sensitivity and Internal AI Hiring

Firms operating in highly regulated, data-sensitive environments face greater risks when outsourcing AI development due to compliance burdens, integration frictions, and the potential for data exposure. Consistent with make-or-buy frameworks, these firms may substitute away from external AI vendors and instead invest in internal AI talent.

Hypothesis 1. *Security-sensitive firms are more likely to invest in internal AI/ML capability building, as proxied by a higher share of AI-related job postings.*

We test this hypothesis using regressions of AI hiring share on security sensitivity, controlling for size (log assets), profitability (ROA), leverage, Tobin's Q, and firm age, and including industry and year fixed effects. The positive association we observe weakens under firm fixed effects, suggesting firm-level heterogeneity in internalization strategies.

3.3. Hypothesis 2a: Supply-Side — Investor Sensitivity and Startup Exits

If security-sensitive firms internalize AI development, their affiliated VC arms may face greater challenges in adopting external AI solutions. Startups backed by these investors may struggle to achieve exits because enterprise adoption is a crucial pathway to acquisition or IPO.

Hypothesis 2a. *AI/ML startups backed by security-sensitive investors are less likely to exit via M&A or IPO, reflecting lower external AI adoption.*

We estimate logistic models of exit (IPO or M&A) on an indicator for whether any investor is security-sensitive, controlling for startup age, total funding, sector, geography, and deal year. We also explore heterogeneity by acquirer type and investor ownership concentration.

3.4. Hypothesis 2b: Supply-Side — Cybersecurity Startups and Strategic Fit

Security-sensitive firms may be reluctant to adopt most external AI solutions, but cybersecurity-focused AI startups—whose products directly address data-risk and compliance—may represent a strategic fit. These firms may therefore experience higher exit likelihood when interacting with sensitive investors or acquirers.

Hypothesis 2b. *Cybersecurity-oriented AI startups are more likely to exit, particularly when backed by or acquired by security-sensitive firms.*

We identify cybersecurity startups using PitchBook vertical classifications and keyword filtering, and estimate interaction models between cybersecurity status and investor sensitivity. With the full 2010–2024 PitchBook sample, we observe the evolution of AI/ML startup exits across multiple regulatory and technological regimes, including the post-2016 wave of enterprise AI deployment. Although our design does not rely on specific regulatory shocks such as GDPR or CCPA, the combination of cross-sectional and time-varying variation in investor sensitivity and acquirer characteristics provides substantial identifying power for studying how data-governance constraints affect external AI adoption.

4. Empirical Framework

This section outlines the empirical framework used to study how data sensitivity shapes firms' internal AI investment (demand side) and the exit outcomes of AI/ML startups (supply side). We begin by defining our classification of security-sensitive firms, investors, and cybersecurity-focused startups—key constructs that underpin all identification strategies. We then describe the empirical models corresponding to Hypotheses 1, 2a, and 2b, followed by robustness and identification considerations.

4.1. Constructing Security-Sensitive Indicators

A central construct in our analysis is the notion of "security-sensitive" actors—firms or investors operating under elevated data protection, privacy, or cybersecurity constraints. These classifications play a central role in our hypotheses: in Hypothesis 1, we identify security-sensitive firms as demand-side adopters; in Hypotheses 2a and 2b, we use investor- and acquirer-side classifications to define exposure on the supply side.

4.1..1 Security-Sensitive Firms (Hypothesis 1)

We classify public firms as security-sensitive using a two-step method: (1) industry codes and (2) business descriptions.

Industry Codes (SIC): We flag industries with high regulatory burdens and data confidentiality exposure:

- *Healthcare (SIC 8000–8099):* HIPAA-covered entities
- *Financial Services (SIC 6000–6799):* GLBA/SEC compliance
- *Defense and Aerospace (SIC 3700–3899):* DFARS and national security

- *Government Contractors (SIC 9100–9199):* Handling federal data
- *Education (SIC 8200–8299):* FERPA-covered institutions

Business Description (Text-Based): We supplement with keyword-based classification using Compustat firm descriptions. Firms are flagged if their descriptions contain terms like "HIPAA," "government contract," "proprietary data," "confidential," "national security," or "classified information." If a firm is flagged via text but not SIC, we manually review it.

Robustness Variants:

- *Conservative:* Must meet both SIC and keyword criteria
- *Expanded:* Satisfy either SIC or keyword rule

All Hypothesis 1 analyses draw on the full Lightcast job postings data available for the period 2010–2024. Results are robust across conservative and expanded versions of the security-sensitivity classification.

4.1..2 Security-Sensitive Investors (Hypotheses 2a-b)

For startup-side analysis, we define security-sensitive exposure at the investor level.

Security-Sensitive Investors: We classify an investor as "sensitive" if it meets any of the following:

- Affiliated with or owned by a firm in the above SIC sectors (e.g., defense contractors, health systems)
- Explicitly linked to a corporate parent with security-sensitive operations
- Government-affiliated (e.g., In-Q-Tel, DARPA partnerships)

We construct a SensitiveBacker dummy at the company-deal-investor level, where the dummy equals 1 if any investor in the deal is sensitive. Then we collapse to the company-deal level using max(SensitiveBacker), which equals 1 if the startup received investment from any sensitive investor in any round.

4.1..3 Cybersecurity Startups (Hypothesis 2b)

We use PitchBook vertical tags and keyword filters to classify startups focused on cybersecurity. Keywords include "encryption," "threat detection," "data protection," "penetration testing," and "cyber infrastructure." We verify these classifications using business descriptions. Startups flagged as cybersecurity-focused are interacted with SensitiveBacker in our analysis.

All Hypotheses 2a-b analyses span the period 2010–2024.

4.2. Empirical Models

4.2..1 Internal AI Hiring (Hypothesis 1)

To test whether security-sensitive firms are more likely to hire internal AI talent, we estimate the following firm-year panel regression:

$$\text{AI Share}_{it} = \beta_1 \cdot \text{Security Sensitive}_i + \beta_2 \cdot X_{it} + \gamma_t + \mu_{st} + \epsilon_{it} \tag{1}$$

Here, AI Share$_{it}$ is the share of AI-related job postings out of total job postings for firm i in year t. The key explanatory variable Security Sensitive$_i$ is a dummy equal to 1 if firm i operates in a security-sensitive industry, as defined using both SIC codes and business description keyword filters (see Section 4.1). Control variables X_{it} include log total assets,

ROA, leverage, Tobin's Q, and firm age. We include year fixed effects γ_t and industry $\times$ year fixed effects μ_{st}, clustering standard errors at the firm level.

Robustness checks include alternative outcomes (e.g., log AI job count, raw AI job count) and placebo regressions using non-AI hiring share. We also conduct subsample analyses before and after 2021, the year in which the U.S. issued Executive Order 14028 on Improving the Nation's Cybersecurity. This order strengthened cybersecurity standards for federal contractors and firms handling sensitive data, especially in healthcare, defense, and financial services. The heightened regulatory scrutiny following the 2021 directive provides a natural setting in which to examine whether security-sensitive firms adjust their internal AI hiring disproportionately.

4.2..2 Exit Probability (Hypothesis 2a)

To test whether startups backed by security-sensitive investors are more likely to exit, we estimate the following logistic regression at the company-deal-year level:

$$\text{Exit}_{ijt} = \beta_1 \cdot \text{Sensitive Backer}_{ijt} + \beta_2 \cdot Z_{ijt} + \theta_t + \lambda_v + \varepsilon_{ijt} \tag{2}$$

The dependent variable Exit_{ijt} equals 1 if startup j experiences an IPO or M&A in year t. $\text{Sensitive Backer}_{ijt}$ is a binary variable equal to 1 if any investor participating in the deal is classified as a security-sensitive acquirer, based on SIC codes and keyword matches. Z_{ijt} includes controls for startup age, total capital raised, deal size, and investor count. We include year fixed effects θ_t and vertical fixed effects λ_v, and cluster standard errors at the startup level. A negative β_1 supports Hypothesis 2a, indicating that startups backed by security-sensitive investors are less likely to exit. This finding is consistent with the interpretation that security-sensitive firms are more cautious about adopting external AI

solutions, preferring to rely on internal development due to data governance and regulatory concerns.

4.2..3 Cybersecurity Focus and Investor Sensitivity (Hypothesis 2b)

To test whether cybersecurity-specialized startups backed by sensitive investors are more likely to exit, we estimate an extended logistic regression with interaction terms:

$$\begin{aligned} \text{Exit}_{ijt} &= \beta_1 \cdot \text{Sensitive Backer}_{ijt} + \beta_2 \cdot \text{Cybersecurity}_j \\ &\quad + \beta_3 \cdot (\text{Sensitive Backer}_{ijt} \times \text{Cybersecurity}_j) + \beta_4 \cdot Z_{ijt} + \theta_t + \lambda_v + \varepsilon_{ijt} \end{aligned} \tag{3}$$

Cybersecurity_j is a dummy equal to 1 if startup j is flagged as cybersecurity-focused, based on vertical classification and keyword parsing. The coefficient β_3 captures the interaction effect, whether cybersecurity startups are more likely to exit when backed by sensitive investors.

4.2..4 Triple Interaction: Post-2021 Cybersecurity Executive Order (Hypothesis 2b)

To exploit exogenous variation from the 2021 Cybersecurity Executive Order, we estimate the following triple-difference model:

$$\text{Exit}_{ijt} = \beta_1 \cdot (\text{Cybersecurity}_j \times \text{Sensitive Backer}_{ijt} \times \text{Post 2021}_t) + \beta_2 \cdot W_{ijt} + \theta_t + \lambda_v + \varepsilon_{ijt} \tag{4}$$

Here, Post 2021_t is a dummy equal to 1 for years 2022 and onward. The Cybersecurity Executive Order, disclosed in mid-2021, significantly elevated regulatory awareness around cybersecurity. If sensitive acquirers became more active in acquiring cybersecurity startups

after the breach, we expect $\beta_1 > 0$. All regressions cluster standard errors at the firm or startup level, depending on unit of observation.

4.2..5 Identification and Robustness

To address endogeneity concerns, we implement several design features and robustness strategies. For Hypothesis 1, forward-looking hiring behavior reduces concerns of reverse causality. For Hypotheses 2a and 2b, we include comprehensive controls for startup characteristics (e.g., age, capital raised, investor count), year fixed effects, and vertical fixed effects.

Sensitive backers are defined using SIC codes and keyword filters, while cybersecurity startups are identified via PitchBook verticals and natural language descriptions. We conduct robustness tests including matched samples, placebo regressions, and alternative exit definitions. Consistent findings across these approaches support the theory that data sensitivity shapes both firm-level internal investment and market-level acquisition strategies in AI/ML innovation ecosystems.

5. Main Results

This section presents the empirical results testing our three main hypotheses. We begin by examining the demand-side behavior of security-sensitive firms (Hypothesis 1), followed by the supply-side dynamics of AI/ML startups (Hypotheses 2a and 2b). Across all models, standard errors are clustered at the firm level (gvkey or CompanyID), and we include appropriate fixed effects to absorb unobserved heterogeneity across time and sectors.

5.1. Hypothesis 1: Security-Sensitive Firms and Internal AI Hiring

We begin by comparing the characteristics of firms across security-sensitive and non-sensitive sectors to establish baseline differences in AI investment behavior. Table 1 reports descriptive statistics for AI hiring metrics and key financial variables. On average, security-sensitive firms exhibit a higher AI job share (mean = 0.023 vs. 0.014) and more AI-related job postings in logs, while also displaying stronger profitability and lower leverage. These patterns suggest a baseline difference in AI investment strategies across firm types.

[Insert Table 1 approximately here.]

We also report the most common AI-related job titles used to construct our dependent variable. Appendix Table A.1 lists the top 15 AI-related job titles identified in the Lightcast data. These include roles such as "machine learning engineer" and "data scientist," which closely align with internal AI capability development.

[Insert Table A.1 approximately here.]

We next visualize trends in AI adoption by sensitivity status. Figure 1 plots the evolution of average AI hiring share from 2010 to 2024. The gap between sensitive and non-sensitive firms grows over time, particularly after the mid-2010s, indicating heightened concern with internal data governance in recent years.

[Insert Figure 1 approximately here.]

We next estimate panel regressions to test whether the descriptive differences in AI hiring persist after controlling for firm characteristics and fixed effects. Table 2 reports the results across four specifications using AI job share, the log of AI job counts, and a placebo outcome (non-AI job share).

[Insert Table 2 approximately here.]

Across specifications, the coefficient on *Security-sensitive* is economically meaningful and statistically significant, though its sign varies with the fixed-effects structure. In Column (1), which includes only year fixed effects, security-sensitive firms exhibit a higher AI job share: the coefficient is 0.010 (s.e. = 0.003), significant at the 1% level. This indicates that firms operating in healthcare, finance, defense, and other sensitive sectors post roughly 1 percentage point higher AI hiring share than other firms—a sizable difference given the average AI job share is around 1–2%.

Column (2) introduces industry $\times$ year fixed effects, which absorb industry-specific time trends in AI adoption. Because security-sensitive sectors (e.g., healthcare, financial services) experienced disproportionately rapid early adoption of AI roles, adding these granular fixed effects mechanically flips the sign to –0.019 (s.e. = 0.010). This reversal reflects the very limited within-industry-year variation in sensitivity classification, rather than a substantive change in economic interpretation. The overall pattern remains consistent with Column (1): security-sensitive sectors have systematically different trajectories of AI hiring.

Column (3) uses the log of AI job counts as the dependent variable. The coefficient on *Security-sensitive* is strongly positive (0.496, s.e. = 0.109), indicating that sensitive firms post substantially more AI-related job openings in levels—a result fully consistent with Hypothesis 1.

As a falsification test, Column (4) replaces the dependent variable with the non-AI job share. The coefficient on Security-sensitive is –0.010 (s.e. = 0.003), indicating a small but statistically precise decrease in the fraction of non-AI roles among sensitive firms. Importantly, this effect runs in the opposite direction of the AI-specific patterns documented in Columns (1)–(3), reinforcing that security-sensitive firms selectively expand AI-related hiring rather than adjusting overall hiring levels. The placebo test thus supports

the interpretation that our main results capture targeted investment in AI capabilities, not broad changes in workforce composition.

Taken together, these results consistently indicate that firms in security-sensitive industries invest more heavily in internal AI/ML talent—either by devoting a greater fraction of their hiring to AI roles or by posting more AI jobs in absolute terms. This pattern aligns closely with Hypothesis 1, which predicts greater internal AI development in environments where data governance risks discourage outsourcing.

5.2. Hypothesis 2a: Exit Outcomes and Security-Sensitive Investors

We next examine whether startups backed by investors from security-sensitive industries are less likely to exit via IPO or M&A. Table 3 reports logistic regressions of exit status on a dummy indicating whether any investor participating in a startup's financing round is classified as security-sensitive.

[Insert Table 3 approximately here.]

Across all specifications, the coefficient on *Sensitive* is negative and highly statistically significant. In Column (1), which includes no fixed effects beyond deal-level controls, the estimated coefficient is –1.111 (s.e. = 0.226). This magnitude implies a substantial reduction in exit likelihood. Converting the logit coefficient into predicted probabilities indicates that, holding deal characteristics constant, startups backed by at least one security-sensitive investor have roughly 60% lower exit probability relative to similar startups backed only by non-sensitive investors. For example, at a baseline exit probability of 15–20%, the corresponding probability for sensitive-backed startups falls to approximately 6–8%.

Columns (2)–(4) incorporate year fixed effects, industry fixed effects, and both sets of fixed effects, respectively. The estimated coefficients remain negative and statistically

significant throughout, ranging from –1.017 to –1.137, demonstrating that the negative association between sensitive investors and exit outcomes is robust to the inclusion of increasingly demanding fixed-effect structures. These results support Hypothesis 2a: investors operating in sectors characterized by heightened data governance and security concerns are less likely to facilitate market exits, consistent with a preference for internalizing AI capabilities rather than acquiring them from external startups.

To complement the regression evidence, Figure 2 plots raw exit rates across four groups defined by investor sensitivity and cybersecurity specialization.

[Insert Figure 2 approximately here.]

The figure reveals a clear pattern: cybersecurity-focused startups backed by sensitive investors exhibit the highest exit rates, followed by cybersecurity startups backed by non-sensitive investors. In contrast, startups that are not cybersecurity-focused show markedly lower exit rates, particularly those backed by sensitive investors. This ordering is informative. While sensitive investors generally depress overall exit probability (consistent with Hypothesis 2a), they are substantially more willing to support exits when a startup's technology directly addresses pressing data security or privacy concerns.

This heterogeneity naturally motivates Hypothesis 2b. The strong exit performance of cybersecurity-focused startups—especially those backed by sensitive investors—suggests that sensitive acquirers selectively pursue external AI/ML solutions in domains that align closely with their own data-protection constraints. We formally test this mechanism in the next subsection using interaction regressions between cybersecurity specialization and investor sensitivity.

5.3. Hypothesis 2b: Cybersecurity Startups and Premium Exits

We now turn to Hypothesis 2b, which examines whether cybersecurity-focused startups—those addressing core data-protection and cyber-risk concerns—are more likely to exit, particularly when backed by investors operating in security-sensitive environments. Table 4 presents logistic regressions of exit likelihood on a cybersecurity indicator and its interaction with investor sensitivity.

[Insert Table 4 approximately here.]

Across all specifications, the coefficient on *Cybersecurity* is positive and statistically significant. In Column (1), cybersecurity-oriented startups are 0.260 more likely (in log-odds terms) to exit compared to other AI startups, consistent with the idea that cybersecurity solutions directly address core data governance constraints faced by potential acquirers. Columns (2)–(4) show similar magnitudes after including year and industry fixed effects.

The coefficient on *Sensitive* remains strongly negative (around –1.1), consistent with Hypothesis 2a, but the interaction term *Cybersecurity* $\times$ *Sensitive* is positive, economically large, and marginally significant (around 2). This implies that although sensitive investors generally depress exit likelihood, they are substantially more willing to facilitate exits when the startup is cybersecurity-focused. In other words, cybersecurity startups overcome the adoption frictions documented in Hypothesis 2a, performing particularly well when their technology aligns with investor data-security needs.

To further illustrate these patterns, Figure 3 plots the share of AI startup acquisitions made by sensitive investors from 2010–2024, separately for cybersecurity startups and all AI startups.

[Insert Figure 3 approximately here.]

The figure shows a clear divergence: while the share of overall AI acquisitions by sensitive investors remains low and relatively flat, acquisitions of cybersecurity startups by these investors steadily increase over time. This selective acquisition pattern is consistent with the regression evidence in Table 4: sensitive investors avoid most external AI solutions but actively acquire cybersecurity startups whose products directly mitigate salient regulatory and data-governance risks.

To probe the mechanism further, we leverage a specific regulatory shock: the 2021 U.S. Executive Order on Improving the Nation's Cybersecurity (EO 14028). The order, issued in mid-2021, heightened federal requirements related to software supply-chain security, Zero Trust architecture, and protection of sensitive data. Because many sensitive investors—including healthcare systems, defense contractors, financial institutions, and government-affiliated entities—operate under strict compliance regimes, EO 14028 plausibly intensified their concerns over external AI/ML solutions and altered acquisition patterns beginning in 2022.

Although the 2021 Cybersecurity Executive Order (EO 14028) is not exogenous in a strict causal sense, it represents a clear and well-timed regulatory tightening that sharply elevated cybersecurity requirements for firms operating in sensitive sectors. We use this policy announcement as a temporal benchmark to examine whether sensitive investors adjusted their acquisition behavior following the increase in regulatory scrutiny.

This regulatory shift enables a difference-in-differences evaluation of how sensitive investors adjusted acquisition behavior relative to non-sensitive investors.

Figure 5 documents event-time estimates of exit probability differences between startups backed by sensitive versus non-sensitive investors in the years surrounding the Executive Order.

[Insert Figure 5 approximately here.]

Pre-order coefficients (years –3 to –1) are small and statistically indistinguishable from zero, supporting the parallel trends assumption. Beginning immediately after EO 14028 (year +1), exit likelihood for sensitive-backed startups declines modestly, with point estimates remaining negative or near zero in the following years. This pattern indicates heightened investor caution or slower acquisition activity post-2021—consistent with increased regulatory scrutiny.

Notably, this post-order decline applies predominantly to non-cybersecurity startups. As we show below, cybersecurity startups experience the opposite pattern when interacting with sensitive investors.

Figure 4 plots the corresponding event study for internal AI hiring among security-sensitive public firms.

[Insert Figure 4 approximately here.]

AI hiring among sensitive firms increases modestly after 2021, with the year +1 coefficient being positive and statistically significant. This mirrors the observed decline in external exits for sensitive-backed startups, reinforcing a central mechanism of this paper: data-sensitive firms increasingly internalize AI capabilities in response to heightened regulatory pressure, reducing willingness to acquire generic external AI solutions while continuing to acquire cybersecurity talent and technologies that are mission-aligned.

6. Robustness and Placebo Tests

We conduct a range of robustness checks and placebo tests to validate our main findings and ensure they are not driven by confounding factors or specification choices.

6.1. Alternative Measures and Placebo Outcomes

To ensure that our results are not driven by scale differences in overall hiring, we re-estimate the main specifications using the log count of AI job postings as the dependent variable. The results are highly consistent with our baseline findings, indicating that security-sensitive firms increase AI-related hiring in both relative (share-based) and absolute terms.

As a placebo test, we replace the dependent variable with the share of non-AI job postings. The coefficient on the security-sensitivity indicator is small and economically insignificant, suggesting that the main effects are specific to AI-related hiring rather than reflecting broad shifts in labor demand.

6.2. Time Trends and Sensitivity Definitions

We examine whether the effect of security sensitivity on AI hiring varies over time by interacting the sensitivity indicator with a post-2021 dummy. The interaction term is positive and statistically significant, consistent with rising organizational attention to data governance after 2020.

We also assess robustness to alternative definitions of security sensitivity. Using a conservative measure that requires both SIC-based classification and keyword matches, as well as a broader version requiring only one of the two, we find stable and consistent results. This confirms that our classification captures meaningful variation in firms' exposure to data-security constraints.

6.3. Clustering, Matching, and Endogeneity

Our results are robust to a variety of inference procedures, including clustering standard errors at the firm, industry, and geographic levels, as well as wild bootstrap clustering.

To further address concerns about selection into security-sensitive sectors, we perform propensity score matching based on firm size, ROA, leverage, and Tobin's Q. Matched-sample estimates continue to show higher AI hiring by security-sensitive firms, indicating that observable firm differences do not drive the results.

Finally, lagged specifications—using one- and two-year lags of the security-sensitivity indicator—yield similar estimates, supporting the temporal ordering that data sensitivity precedes subsequent increases in internal AI hiring.

6.4. Startup-Side Placebo Tests

For the startup-side analyses (Hypotheses 2a and 2b), we conduct placebo tests focusing on AI/ML startups operating in low-data-sensitivity sectors such as entertainment, media, and e-commerce. In these subsamples, neither the presence of a security-sensitive investor nor cybersecurity specialization is significantly related to exit outcomes. The absence of effects in low-sensitivity domains reinforces our interpretation that the main results are concentrated in environments where data governance, regulatory obligations, and internal development concerns are most salient.

7. Alternative Explanations

While our results support the view that data sensitivity shapes internal AI hiring and startup exit dynamics, several alternative mechanisms could also contribute to the observed patterns. Here, we outline potential explanations and discuss how our empirical design addresses them, as well as additional tests we plan to implement.

A first possibility is that firms in security-sensitive sectors are simply more innovative or technologically advanced. Such firms may naturally exhibit greater AI hiring due to

higher RD intensity or stronger technological capabilities. Our baseline regressions already control for firm size, profitability, leverage, and Tobin's Q. In future robustness checks, we will further incorporate RD expenditures and patenting activity where available to assess whether innovation intensity can account for the observed patterns. Planned placebo regressions on non-AI job shares will help determine whether the effect is specific to AI-related hiring.

Another explanation is that sector-wide labor trends may drive the main results. For example, healthcare and financial services may be increasing data-related hiring due to broader industry automation. Our specifications include two-digit SIC and industry-by-year fixed effects, which already address much of this concern. As an additional robustness step, we will explore whether allowing for finer-grained industry controls alters the estimated relationship between security sensitivity and AI hiring.

Classification concerns also arise from our use of Lightcast job postings. Firms could differ in how they label technical positions, potentially generating measurement error in AI job counts. To evaluate this, we intend to replicate the analysis using alternative AI-related keyword filters, as well as occupation-level taxonomies, including the log count of AI jobs. Consistency across these alternative measures would reduce concerns about classification bias.

Labor supply differences provide another potential explanation. If security-sensitive firms are disproportionately located in regions with greater AI talent availability, they may hire more AI workers regardless of internal data-governance concerns. Our baseline analysis includes state fixed effects to account for regional variation. As a follow-up test, we plan to incorporate more granular geographic controls (e.g., commuting-zone or metropolitan-area fixed effects) and test for differential results in non-AI technical roles.

Another possibility is that firms in sensitive sectors may hire AI talent for customer-facing applications rather than internalization. To investigate this, we will examine the

functional breakdown of AI job postings (e.g., infrastructure, compliance, customer-facing automation). If sensitive firms disproportionately hire for back-end machine learning, cybersecurity, or data-governance roles, this would support our interpretation.

Finally, reverse-causality concerns arise if firms alter their descriptions or reporting in ways correlated with AI adoption. Our classification of security sensitivity is based on SIC codes and business descriptions that are largely time-invariant and determined prior to AI hiring decisions. Planned lagged-regression tests will assess whether security sensitivity predicts future AI hiring rather than the reverse.

In sum, although several alternative explanations—innovation intensity, sector-wide trends, measurement error, labor supply, or customer-facing automation—may influence hiring decisions, our current identification strategy already addresses many of these channels. Additional robustness tests, outlined above, will help clarify the extent to which these alternatives can account for the strong and specific relationship we document between data sensitivity and AI adoption on the demand side, and between investor sensitivity and exit outcomes on the supply side.

8. Discussion and Implications

Our findings offer important insights into how data sensitivity shapes the adoption and commercialization of AI/ML technologies. The evidence from both acquirer firms (demand side) and service provider startups (supply side) highlights the critical role of data governance in mediating technological innovation and strategic decision-making. In this section, we interpret the broader significance of our results and outline several implications for theory, corporate strategy, startup dynamics, and public policy.

8.1. Theoretical Implications: Innovation under Institutional Constraints

Our findings contribute to the broader literature on innovation under institutional and regulatory frictions. While prior work on AI adoption often emphasizes technological readiness, resource endowments, or human capital (Brynjolfsson and McElheran (2016); Babina et al. (2024)), the evidence presented here underscores a complementary and often underappreciated force: firms' exposure to data governance constraints. Security-sensitive firms internalize AI not simply because they possess superior capabilities or a stronger innovation orientation, but because the risks associated with outsourcing—data leakage, privacy violations, reputational harm—are difficult to contract upon.

This mechanism connects directly to theories of transaction costs and incomplete contracting (Williamson (1985); Grossman and Hart (1986)). When key contingencies cannot be specified ex ante, and when the cost of a breach is asymmetric or extremely high, internalizing AI capability becomes a way to mitigate non-contractible risks. Our results suggest that regulatory exposure and institutional constraints can act as substitutes for traditional determinants of firm boundaries, motivating internal development in domains where outsourcing may be otherwise efficient. This provides empirical grounding for recent theoretical arguments that institutional risk increasingly shapes innovation strategy in data-intensive environments.

8.2. Implications for Incumbent Firms: Strategic AI Sourcing

For incumbent firms, particularly those in healthcare, finance, and defense, our findings imply that sourcing AI is fundamentally a risk management decision. Rather than purchasing off-the-shelf AI products or relying on external vendors, these firms favor internal AI hiring to maintain tighter control over sensitive data and compliance processes. This approach

enhances data integrity and reduces exposure to external vulnerabilities, but it may also lead to slower implementation cycles, higher fixed costs, and potential duplication of effort across firms.

In practice, the choice between building and buying is shaped not only by technological needs but also by institutional pressures. Firms in highly regulated sectors evaluate AI vendors through a narrower lens that emphasizes data governance and regulatory compliance. As a result, internalization becomes the dominant strategy even when outsourcing could accelerate adoption or leverage specialized external expertise. This insight highlights the strategic tension incumbents face: balancing innovation objectives with the need to protect sensitive information and maintain regulatory alignment.

8.3. Implications for AI/ML Startups: Strategic Positioning and Exit Prospects

For AI/ML startups, the results highlight the importance of strategic alignment with the risk and compliance frameworks of potential acquirers. Startups that articulate their value proposition in ways that resonate with buyers' data governance concerns—such as cybersecurity-focused firms—experience higher exit rates and, in many cases, premium valuations. These startups help reduce perceived regulatory risk and thereby mitigate the internalization bias prevalent among sensitive acquirers.

By contrast, AI products that do not directly address sensitive data concerns face greater adoption friction. These startups may struggle to secure enterprise contracts or achieve liquidity events, even if their technologies are technically superior. The market for exits is therefore segmented: cybersecurity-focused AI firms benefit from strong acquirer demand, while more generic AI/ML startups face a higher bar to integration in sensitive sectors. This dynamic has implications for startup strategy, highlighting the importance of domain

expertise, compliance awareness, and early alignment with industry-specific governance norms.

8.4. Policy Implications: Regulation and Innovation Alignment

The results also carry implications for policymakers seeking to balance data protection with innovation diffusion. Fragmented or overly stringent data regulations can inadvertently encourage firms to internalize AI development, reducing specialization and slowing the spread of frontier technologies. Internalization can be costly, duplicative, and inefficient, particularly when firms independently rebuild similar AI capabilities.

However, the findings also point to actionable pathways for policy design. Clearer standards for data handling, third-party audits, and vendor compliance—such as certification schemes or safe-harbor provisions—may facilitate trusted outsourcing and reduce the need for firms to rebuild capabilities in-house. Regulatory clarity can help shift the make-or-buy calculus back toward efficient specialization and foster a more competitive and interoperable ecosystem for AI innovation.

In summary, the evidence suggests that data governance constraints shape both internal adoption and external acquisition in the AI ecosystem. The interplay between institutional risk, firm boundaries, and innovation sourcing has important consequences not only for the distribution of AI capabilities across firms but also for the competitive dynamics between incumbents and startups.

9. Conclusion

This paper examines how data governance concerns shape both the internalization of AI capabilities by incumbent firms and the commercialization pathways of AI/ML startups.

Using detailed job posting data from Lightcast combined with firm-level financial information from Compustat, we show that firms operating in security-sensitive industries—such as healthcare, finance, and defense—are significantly more likely to build internal AI capacity through specialized hiring. These patterns suggest that internalization is not merely a reflection of technological readiness or resource advantages, but rather a strategic response to regulatory exposure, data-protection obligations, and concerns over third-party vulnerability.

On the supply side, we investigate how these same institutional frictions influence the exit prospects of AI/ML startups. Cybersecurity-focused startups—those whose products directly mitigate data-risk concerns—experience substantially higher exit rates, particularly in periods marked by heightened regulatory attention. These firms appear to overcome the adoption frictions that limit exits for more general-purpose AI vendors, indicating a strong alignment between incumbent demand and specialized startup capabilities.

Taken together, our results highlight the broader role of institutional constraints in shaping the boundaries of AI innovation. Theoretically, they enrich models of make-or-buy decisions by demonstrating how data governance risks alter the tradeoff between internal development and external sourcing. Managerially, they underscore the importance of internal hiring as an essential component of risk management in AI adoption. For startups, the findings emphasize the value of signaling compliance readiness, domain-specific credibility, and alignment with the regulatory expectations of potential acquirers. From a policy perspective, our results suggest that regulatory clarity—not only regulatory stringency—can facilitate more efficient diffusion of AI technologies by enabling trusted engagement between incumbents and specialized vendors.

10.1. Future Research: Regulatory Nuance and Consistency in AI Startup Positioning

While our primary analysis focuses on internal AI hiring and binary exit outcomes, future research can explore the subtler ways in which regulatory nuance shapes heterogeneity in commercialization success. Appendix Figures A.7 to A.8 and Table A.3 introduce a "consistency score" that measures the similarity between a startup's business description and the regulatory vocabulary commonly used by incumbents in sensitive industries.

Preliminary evidence suggests that startups with higher consistency scores—those that more closely mirror the linguistic framing and compliance narratives of their target sectors—exhibit higher exit probabilities, especially when backed by security-sensitive investors. This highlights a more granular channel operating alongside domain specialization: beyond offering cybersecurity functionality, the way startups articulate their value proposition and regulatory fit may influence acquirer perceptions and lower integration frictions.

These findings open promising avenues for further work. Future research could incorporate NLP-based measures of regulatory mimicry using regulatory filings, certifications, audit disclosures, or security frameworks to capture startups' alignment with data-governance expectations. Such extensions would deepen our understanding of how firms navigate institutional frictions and may help policymakers and investors identify startups that are not only technologically strong but also institutionally credible in data-sensitive environments.

References

Abis, S. (2020). Man vs. Machine: Quantitative and Discretionary Equity Management. *Working Paper*.

Abis, S. and L. Veldkamp (2024). The Changing Economics of Knowledge Production. *Review of Financial Studies 37*(1), 89–118.

Acemoglu, D., D. Autor, J. Hazell, and P. Restrepo (2022). Artificial Intelligence and Jobs: Evidence from Online Vacancies. *Journal of Labor Economics 40*(S1).

Alekseeva, L., M. Gine, S. Samila, and B. Taska (2020). AI Adoption and Firm Performance: Management versus IT. *Working Paper 96*(4), 1069–1090.

Anderson, R. and T. Moore (2006). The Economics of Information Security. *Science 314*(5799), 610–613.

Arora, A. and A. Gambardella (1990). Complementarity and External Linkages: The Strategies of the Large Firms in Biotechnology. *Journal of Industrial Economics 38*(4), 361–379.

Babina, T., S. Bahaj, G. Buchak, F. De Marco, A. Foulis, W. Gornall, F. Mazzola, and T. Yu (2025). Customer data access and fintech entry: Early evidence from open banking. *Journal of Financial Economics 169*(103950).

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

Bach, L., R. P. Baghai, P. Strömberg, and K. Warg (2022). Who becomes a business angel? *Working Paper*.

Barrot, J.-N. (2017). Investor Horizon and the Life Cycle of Innovative Firms: Evidence from Venture Capital. *Working Paper*.

Beggs, W., J. Brogaard, and A. Hill-Kleespie (2025). Quantitative Investing and Market Instability: Evidence from Mutual Fund Fire Sales. *Working Paper*.

Bellardini, L., B. L. Del Gaudio, D. Previtali, and V. Verdoliva (2022). How do banks invest in fintechs? Evidence from advanced economies. *Journal of International Financial Markets, Institutions and Money 77*(101498).

Bena, J. and K. Li (2014). Corporate Innovations and Mergers and Acquisitions. *Journal of Finance 69*(5), 1923–1960.

Bereskin, F., S. K. Byun, M. S. Officer, and J.-M. Oh (2018). The Effect of Cultural Similarity on Mergers and Acquisitions: Evidence from Corporate Social Responsibility. *Journal of Financial and Quantitative Analysis 53*(5), 1995–2039.

Berk, J. B. and R. C. Green (2004). Mutual Fund Flows and Performance in Rational Markets. *Journal of Political Economy 112*(6), 1269–1295.

Bertomeu, J., Y. Lin, Y. Liu, and Z. Ni (2023). On the Value Losses of Disruption in Technology Adoption. *Working Paper*.

Bezaei, R. and Y. Yao (2022). Venture Capital Response to Government-Funded Basic Science. *Working Paper*.

Bloom, N., E. Brynjolfsson, L. Foster, R. Jarmin, M. Patnaik, I. Saporta-Eksten, and J. Van Reenen (2019). What Drives Differences in Management Practices? *American Economic Review 109*(5), 1648–1683.

Bonelli, M. (2022). Data-driven Investors. *Working Paper*.

Bresnahan, T. F., E. Brynjolfsson, and L. M. Hitt (2002). Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence. *Quarterly Journal of Economics 117*(1), 339–376.

Brynjolfsson, E., D. Li, and L. R. Raymond (2023). Generative AI at Work. *NBER Working Paper*.

Brynjolfsson, E. and K. McElheran (2016). The Rapid Adoption of Data-Driven Decision-Making. *American Economic Review 106*(5), 133–139.

Buchak, G., G. Matvos, T. Piskorski, and A. Seru (2018). Fintech, regulartory arbitrage, and the rise of shadow banks. *Journal of Financial Economics 130*(3), 453–483.

Canayaz, M. and Z. Wang (2023). Crafting an AI Compass: The Influence of Global AI Standards on Firms. *Working Paper*.

Cao, S., Y. Cheng, M. Wang, Y. Xia, and B. Yang (2023). Visual Information and AI Divide: Evidence from Corporate Executive Presentations. *Working Paper*.

Casadesus-Masanell, R. and J. E. Ricart (2010). From Strategy to Business Models and onto Tactics. *Long Range Planning 43*(2-3), 195–215.

Cavusoglu, H., B. Mishra, and S. Raghunathan (2004). The Effect of Internet Security Breach Announcements on Market Value: Capital Market Reactions for Breached Firms and Internet Security Developers. *International Journal of Electronic Commerce 9*(1), 69–104.

Cen, X. (2024). Smartphone Trading Technology, Investor Behavior, and Mutual Fund Performance. *Management Science 70*(10), 6483–7343.

Chemmanur, T. J., M. B. Imerman, H. Rajaiya, and Q. Yu (2023). The Entrepreneurial Finance of Fintech Firms and the Effect of Investments in Fintech Startups on the Performance of Corporate Investors. *Working Paper*.

Coase, R. H. (1937). The Nature of the Firm. *Economica 4*(16), 386–405.

Dessaint, O., T. Foucault, and L. Fresard (2024). Does Alternative Data Improve Financial Forecasting? The Horizon Effect. *Journal of Finance 79*(3), 2237–2287.

Di Maggio, M. and V. Yao (2020). Fintech Borrowers: Lax Screening or Cream-Skimming? *Review of Financial Studies 34*(10), 4565–4618.

Ewens, M. and J. Farre-Mensa (2020). The Deregulation of the Private Equity Markets and the Decline in IPOs. *Review of Financial Studies 33*(12), 5463–5509.

Ewens, M., R. Nanda, and M. Rhodes-Kropf (2018). Cost of experimentation and the evolution of venture capital. *Journal of Financial Economics 128*(3), 422–442.

Farboodi, M., A. Matray, L. Veldkamp, and V. Venkateswaran (2021). Where Has All the Data Gone? *Review of Financial Studies 35*(7), 3101–3138.

Gal, M. and D. L. Rubinfeld (2019). Data Standardization. *New York University Law Review 94*(737), 738–770.

Gans, J. S. and S. Stern (2003). The product market and the market for "ideas": commercialization strategies for technology entrepreneurs. *Research Policy 32*(2), 333–350.

Gordon, L. A., M. P. Loeb, W. Lucyshyn, and L. Zhou (2015). Increasing cybersecurity investments in private sector firms. *Journal of Cybersecurity 1*(1), 3–17.

Grennan, J. and R. Michaely (2020). Artificial Intelligence and High-Skilled Work: Evidence from Analysts. *Swiss Finance Institute Research Paper*.

Grossman, S. J. and O. D. Hart (1986). The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration. *Journal of Political Economy 94*(4).

Harvey, C. R., S. Rattray, A. Sinclair, and O. V. Hemert (2017). Man vs. Machine: Comparing Discretionary and Systematic Hedge Fund Performance. *Duke IE Resesarch Paper*.

Hellman, T. (2002). A theory of strategic venture investing. *Journal of Financial Economics 64*(2), 285–314.

Hellmann, T. and M. Puri (2002). Venture Capital and the Professionalization of Start-Up Firms: Empirical Evidence. *Journal of Finance 57*(1), 169–197.

Karlsen, J., K. Kisseleva, A. Mjøs, and D. T. Robinson (2024). Are Some Angels Better than Others? *Working Paper*.

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

Ko, B., J. Sheng, and Z. Sun (2023). Capitalizing on Retail Investor Sentiment: Evidence from FinTech ETFs. *Working paper*.

Li, E., M. Q. Mao, H. F. Zhang, and H. Zheng (2023). Banks' investments in fintech ventures. *Journal of Banking and Finance 149*(106754).

Mao, M. Q. and H. Zheng (2024). Fintech mergers and acquisitions. *Journal of International Money and Finance 143*(103076).

Masclans, R. (2025). Science, Startups, and the Problem of Value Capture: Thin Acquisition Markets, Weak Outside Options. *Working paper*.

Mathews, R. D. (2006). Strategic alliances, equity stakes, and entry deterrence. *Journal of Financial Economics 80*(1), 35–79.

Metrick, A. and A. Yasuda (2007). *Venture Capital and the Finance of Innovation*. Wiley.

Mittal, V. (2024). Desperate Capital Breeds Productivity Loss: Evidence from Public Pension Investments in Private Equity. *Working Paper*.

Paine, F. (2025). Not on Terra FIRRMA: Foreign Investment in US Startups and Innovation. *Working Paper*.

Pham, P. K., R. Rezaei, and J. Zein (2023). Venture Capitalists vs. Deep-Pocketed Incumbents: Startup Financing Strategies in the Presence of Competitive Threats. *UNSW Business School Research Paper Forthcoming*.

Puri, M., Y. Qian, and X. Zheng (2024). From Competitors to Partners: Banks' Venture Investments in Fintech. *Working Paper*.

Pástor, , R. F. Stambaugh, and T. L. A. (2021). Fund trade-offs. *Journal of Financial Economics 142*(2), 550–571.

Romanosky, S. (2016). Examining the costs and causes of cyber incidents. *Journal of Cybersecurity 2*(2), 121–135.

Webb, M. (2019). The Impact of Artificial Intelligence on the Labor Market. *Working Paper*.

Williamson, O. E. (1985). *The Economic Institutions of Capitalism*. Free Press.

Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. PublicAffairs.

10. Tables and Figures

Figure 1. AI Hiring Share Over Time by Security Sensitivity

This figure plots the average share of AI-related job postings among total job postings at the firm-year level, separately for security-sensitive and non-sensitive firms, between 2010 and 2024. Security-sensitive firms include those in industries such as healthcare, defense, and government services, identified by SIC and NAICS codes. The blue line represents security-sensitive firms, while the yellow line shows non-sensitive firms. The average AI hiring share is consistently higher for security-sensitive firms starting mid-2010s.

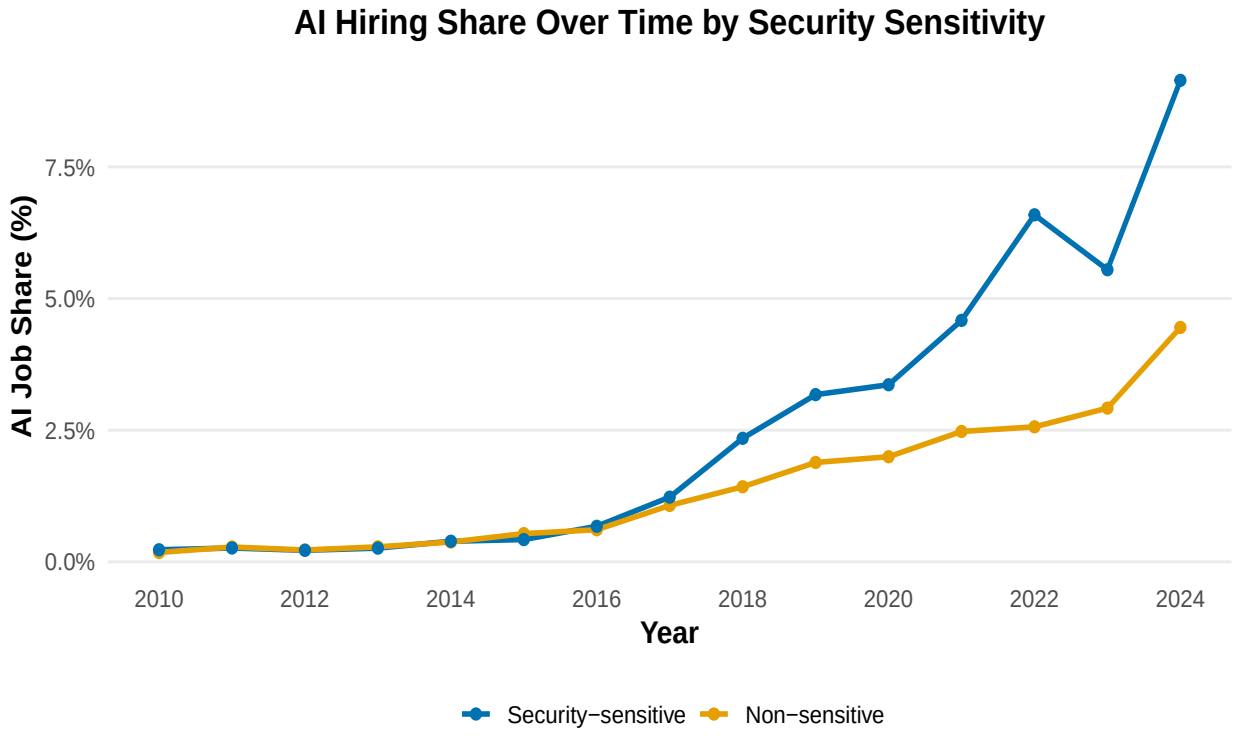

Figure 2. Exit Rates by Cybersecurity Specialization and Investor Sensitivity

This figure presents the raw exit rates for four startup groups based on their cybersecurity specialization and whether they were backed by security-sensitive investors. Cybersecurity startups backed by sensitive investors ("Cyber + Sensitive") exhibit the highest exit rates, consistent with Hypotheses 2a and 2b.

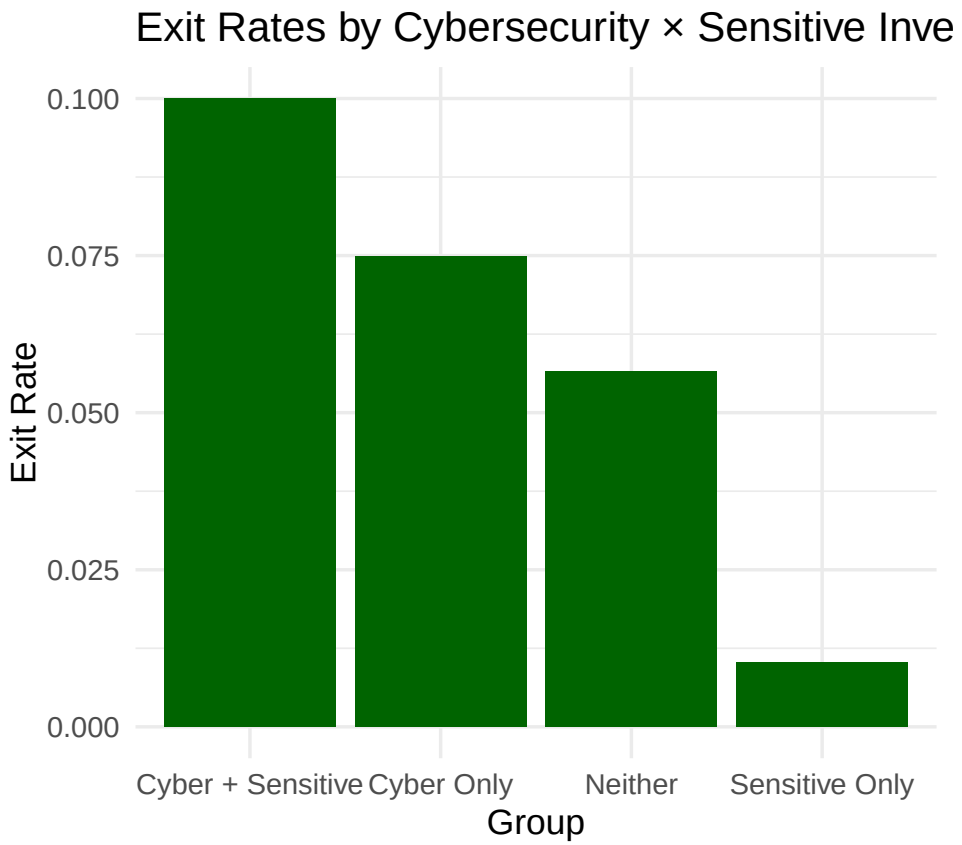

Figure 3. Share of AI Startup Acquisitions by Security-Sensitive Investors (2010–2024)

This figure plots the yearly share of AI startup acquisitions made by security-sensitive investors from 2010 to 2024. The solid line represents all AI startups, while the dashed line highlights cybersecurity startups. Cybersecurity acquisitions by sensitive investors steadily increased, suggesting these investors are more willing to acquire mission-aligned AI startups. In contrast, overall AI acquisitions by sensitive investors remain limited, consistent with internal development or caution toward external AI adoption.

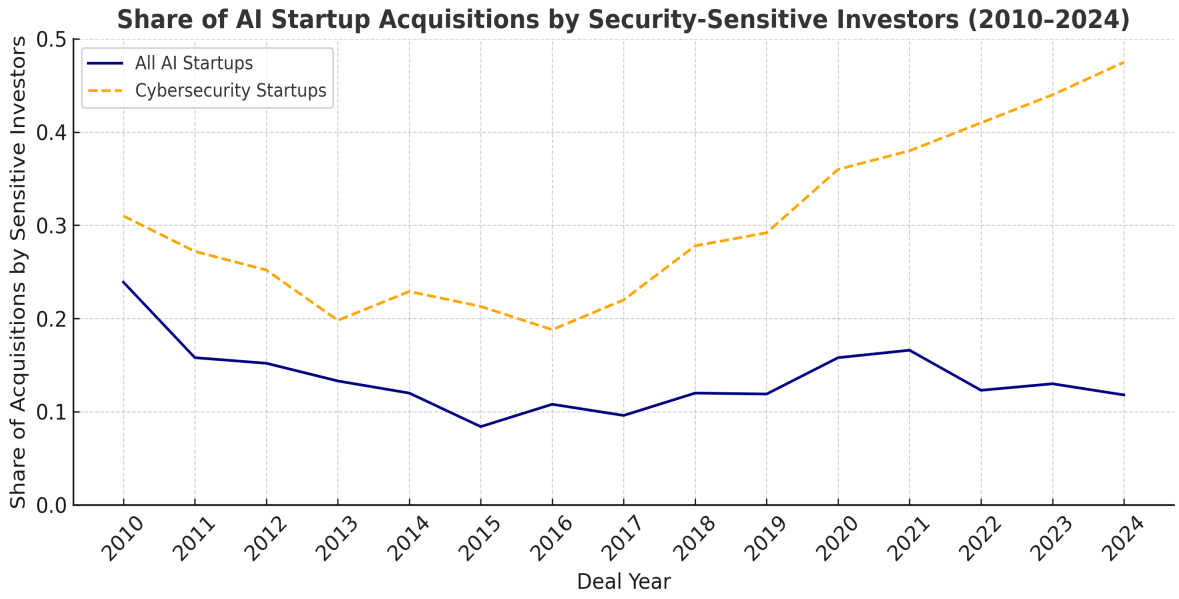

Figure 4. Event Study: AI Hiring After Cybersecurity Executive Order (±3 Years)

This figure presents event-time estimates of AI hiring share around the 2021 Cybersecurity Executive Order (EO 14028). The coefficients represent the difference in AI hiring between security-sensitive and non-sensitive firms in each year relative to 2021 (baseline = year 0). The dashed vertical line at year 0 marks the order. Pre-order coefficients are mildly negative and marginally significant. The year 1 post-order coefficient is positive and significant at the 5% level, while later years are smaller and statistically weaker, indicating a gradual adjustment in AI hiring following the order.

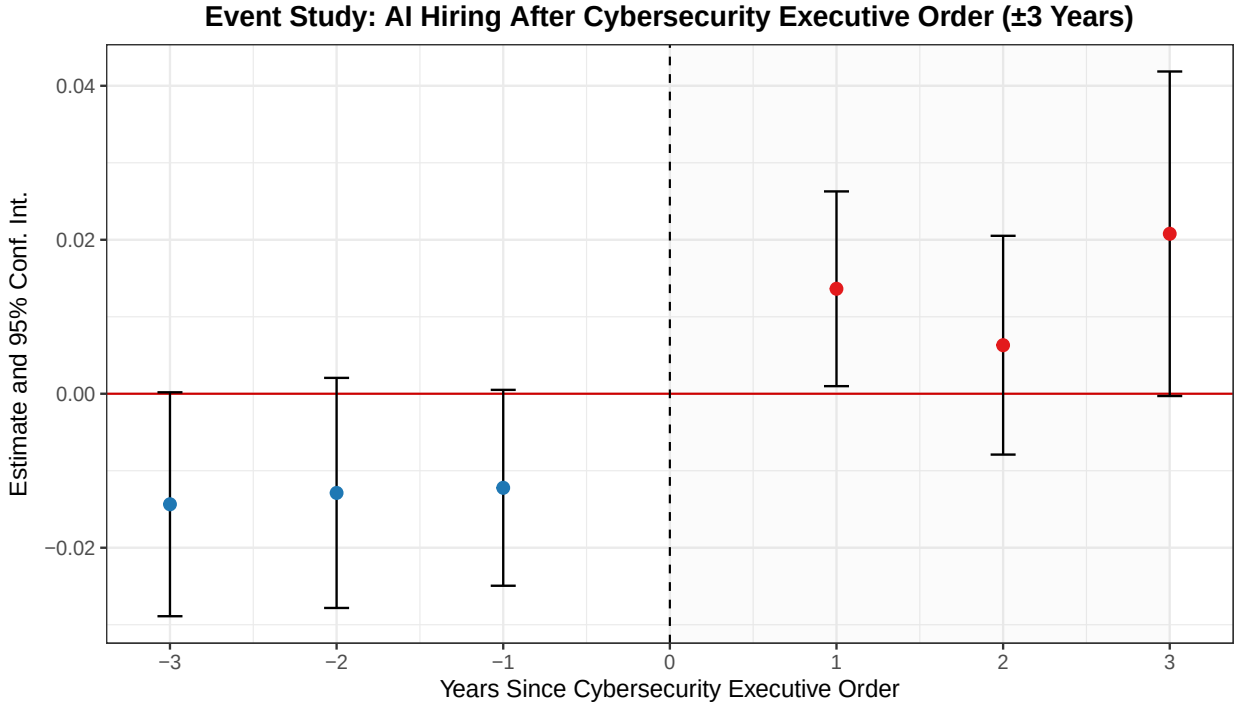

Figure 5. Event Study: Startup Exit Trends Around the 2021 Cybersecurity Executive Order (±3 Years)

This figure presents event-time estimates of startup exit likelihood around the 2021 Cybersecurity Executive Order (EO 14028). The coefficients represent the difference in exit probability between startups backed by sensitive versus non-sensitive investors in each year relative to 2021 (baseline = year 0). The dashed vertical line at year 0 marks the issuance of the order. Pre-order coefficients are relatively small and statistically weak, while the first post-order estimates show modest declines in exit likelihood for startups with sensitive investors, suggesting heightened investor caution and slower acquisition dynamics following the Executive Order.

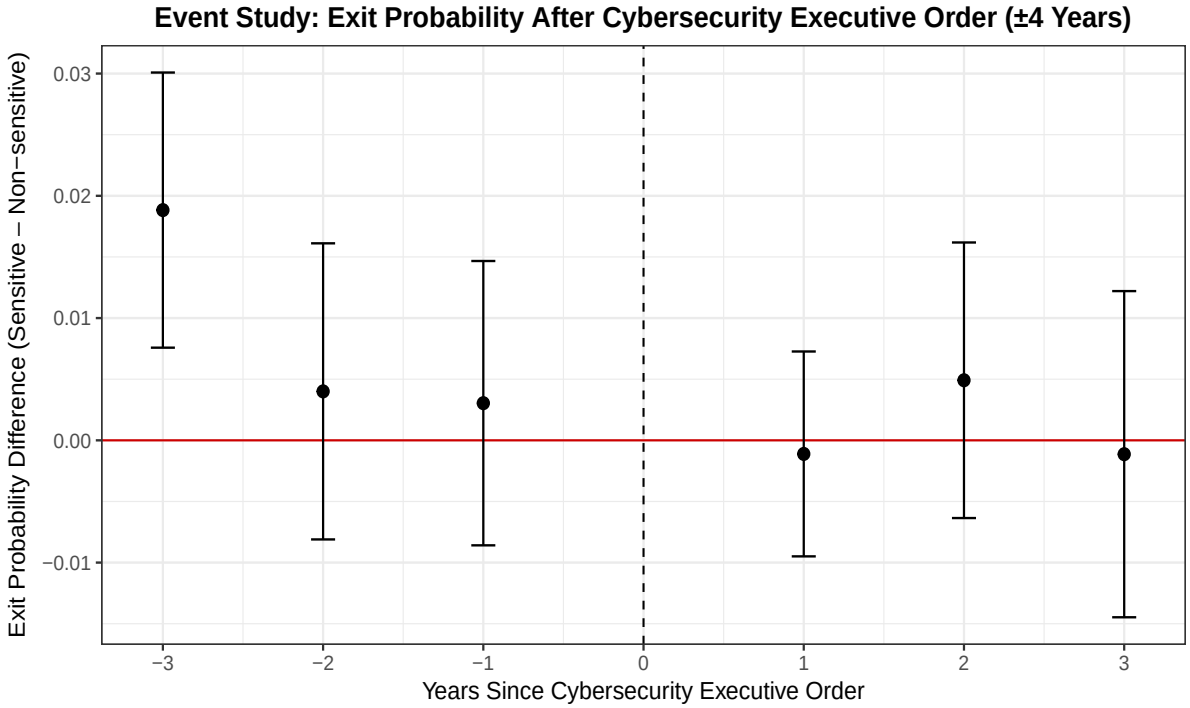

Table 1. Summary Statistics by Security Sensitivity Status

This table presents summary statistics for AI hiring variables and financial indicators among firms classified as security-sensitive (e.g., healthcare, defense) vs. others. The AI job share metric reflects the proportion of AI-tagged postings relative to all matched postings for a firm-year. N varies by variable due to incomplete reporting of firm-level financials and matched job data across firm-years.

Variable	Group	N	Mean	SD	25p	75p
AI Job Share	Not Sensitive	42,033	0.014	0.071	0.000	0.003
	Sensitive	5,129	0.023	0.098	0.000	0.008
Log(AI Jobs)	Not Sensitive	42,033	0.755	1.434	0.000	1.099
	Sensitive	5,129	1.037	1.760	0.000	1.609
Non-AI Job Share	Not Sensitive	42,033	0.986	0.071	0.997	1.000
	Sensitive	5,129	0.977	0.098	0.992	1.000
Log(Assets)	Not Sensitive	41,638	7.280	2.532	5.653	8.929
	Sensitive	5,082	7.031	2.617	5.309	8.834
ROA	Not Sensitive	41,618	-0.080	0.462	-0.041	0.056
	Sensitive	5,081	-0.058	0.494	-0.036	0.073
Leverage	Not Sensitive	41,574	0.679	0.538	0.426	0.819
	Sensitive	5,075	0.650	0.630	0.329	0.783
Tobin's Q	Not Sensitive	35,094	2.329	2.466	1.110	2.490
	Sensitive	4,574	2.405	2.537	1.196	2.554

Table 2. Security-Sensitive Firms and Internal AI Hiring

This table presents the relationship between security-sensitive firm status and internal AI hiring across different model specifications. The key independent variable is *Security-sensitive*, an indicator equal to 1 for firms operating in security-sensitive industries. The dependent variables include the AI job share, log of AI jobs, and a placebo test using non-AI job share. Column (2) includes industry×year fixed effects, which absorb industry-level time trends in AI hiring. Because security-sensitive sectors (e.g., healthcare, finance) have high baseline adoption, the within–industry-year residualized specification mechanically flips the sign, though the overall pattern remains consistent with Columns (1), (3), and (4). All regressions control for firm size (log(at)), ROA (sales/assets), leverage, and Tobin's Q. Standard errors are clustered at the firm level (gvkey).

	Dependent Variable			
	(1) AI Job Share	(2) AI Job Share	(3) Log(AI Jobs)	(4) Non-AI Job Share (Placebo)
Security-sensitive	0.010***	-0.019*	0.496***	-0.010***
	(0.003)	(0.010)	(0.109)	(0.003)
Log(at)	0.001***	0.002***	0.426***	-0.001***
	(0.000)	(0.000)	(0.020)	(0.000)
ROA	-0.002	-0.005**	-0.201*	0.002
	(0.002)	(0.002)	(0.103)	(0.002)
Leverage	-0.007***	-0.003	-0.132*	0.007***
	(0.002)	(0.002)	(0.080)	(0.002)
Tobin's Q	0.003***	0.001*	0.183***	-0.003***
	(0.001)	(0.001)	(0.014)	(0.001)
Fixed Effects				
Year FE	Yes	Yes	Yes	Yes
Industry x Year FE		Yes		
Observations	39,658	39,658	13,173	39,658
R^2	0.045	0.202	0.270	0.045
R^2 (within)	0.011	0.005	0.240	0.011
RMSE	0.08	0.07	1.59	0.08

Standard errors clustered by firm (gvkey); t-stats in parentheses
*Signif. Codes: + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001*

Table 3. Startup Exits and Investor Sensitivity

This table presents results from logistic regressions of startup exit on whether the investor operated in a security-sensitive domain. Controls included are DealSize, RaisedtoDate, and Investors. DealSize is the total amount of capital invested into a company by an investor or group of investors for a specific transaction (in millions), RaisedtoDate is the sum of known capital injected since last majority transaction or recapitalization, and Investors is the count of the total investors in the deal. Fixed effect for DealYear is included. This is at the company-deal level. All models cluster standard errors by startup (CompanyID).

	Dependent Variable: Exit (1 = IPO or Acquisition)			
Variables	(1) No FE	(2) DealYear FE	(3) Industry FE	(4) Both FE
Sensitive	-1.111***	-1.137***	-1.017***	-1.044***
	(0.226)	(0.179)	(0.221)	(0.185)
Fixed Effects and Controls				
DealYear		Yes		Yes
PrimaryIndustryGroup			Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	99,463	99,463	99,380	99,380
R^2	0.056	0.060	0.062	0.066
R^2 (within)		0.056	0.055	0.056
RMSE	0.23	0.23	0.23	0.23

Standard errors clustered by startup (CompanyID); std. errors in parentheses
*Signif. Codes: + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001*

Table 4. Cybersecurity Startups and Exit Likelihood in Sensitive Sectors

This table presents results from logistic regressions of startup exit on cybersecurity specialization and investor sensitivity. The interaction term indicates whether a cybersecurity startup was backed by a sensitive investor. Controls included are DealSize, RaisedtoDate, and Investors. DealSize is the total amount of capital invested into a company by an investor or group of investors for a specific transaction (in millions), RaisedtoDate is the sum of known capital injected since last majority transaction or recapitalization, and Investors is the count of the total investors in the deal. Fixed effects for PrimaryIndustryGroup and DealYear are included. This is at the company-deal level. All models cluster standard errors by startup (CompanyID).

	Dependent Variable: Exit (1 = IPO or Acquisition)			
Variables	(1) No FE	(2) DealYear FE	(3) Industry FE	(4) Both FE
Cybersecurity	0.260**	0.265**	0.197***	0.203*
	(0.089)	(0.091)	(0.032)	(0.087)
Sensitive	-1.140***	-1.165***	-1.046***	-1.073***
	(0.232)	(0.169)	(0.249)	(0.172)
Cybersecurity × Sensitive	2.072+	2.021*	2.032	1.985*
	(1.149)	(0.965)	(1.651)	(0.971)
Fixed Effects and Controls				
DealYear		Yes		Yes
PrimaryIndustryGroup			Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	99,463	99,463	99,380	99,380
R^2	0.056	0.060	0.062	0.066
R^2 (within)		0.056	0.055	0.056
RMSE	0.23	0.23	0.23	0.23

Standard errors clustered by startup (CompanyID); std. errors in parentheses
*Signif. Codes: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$*
Cybersecurity × Sensitive = Interaction of cybersecurity specialization and sensitive investor

Table 5. Difference-in-Differences: Exit Likelihood Around Cybersecurity Executive Order (2021)

This table presents the results of DID (Column 1) and Triple DID (Column 2) regressions estimating the effect of the 2021 Executive Order (EO 14028) on the likelihood of startup exit. The DID specification tests whether exits increased for startups backed by sensitive investors. The Triple DID specification adds a Cybersecurity specialization interaction. All regressions include DealYear and PrimaryIndustryGroup fixed effects and control for DealSize, RaisedtoDate, and Investors. Standard errors are clustered by CompanyID.

	(1) DID: Sensitive Only	(2) Triple DID: Cyber × Sensitive
Cybersecurity		0.009
		(0.029)
Sensitive	-0.044***	-0.044***
	(0.010)	(0.010)
Post × Sensitive	0.037*	0.031*
	(0.015)	(0.014)
Post × Cybersecurity		-0.006
		(0.030)
Cybersecurity × Sensitive		0.849***
		(0.015)
Controls		
log(DealSize)	0.021***	0.021***
	(0.001)	(0.001)
log(RaisedtoDate + 1)	-0.006***	-0.006***
	(0.002)	(0.002)
Investors	-0.007***	-0.007***
	(0.000)	(0.000)
Fixed Effects		
DealYear	Yes	Yes
PrimaryIndustryGroup	Yes	Yes
Observations	40,880	40,880
Pseudo R^2	-0.074	-0.075
RMSE		

Standard errors clustered by CompanyID; std. errors in parentheses
*Signif. Codes: + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001*

Appendix

Figure A.1. AI job share over time

This figure shows the share of U.S. job postings that mention AI-related roles or skills, using the full Lightcast (Burning Glass) dataset from 2010 through May 2025. The series reflects the fraction of postings requiring machine learning, artificial intelligence, data science, or related competencies. AI hiring remained relatively modest until the mid-2010s, followed by a sustained increase and a sharp acceleration beginning in 2023. The trend highlights the broad and rapid expansion of internal AI capability building across firms.

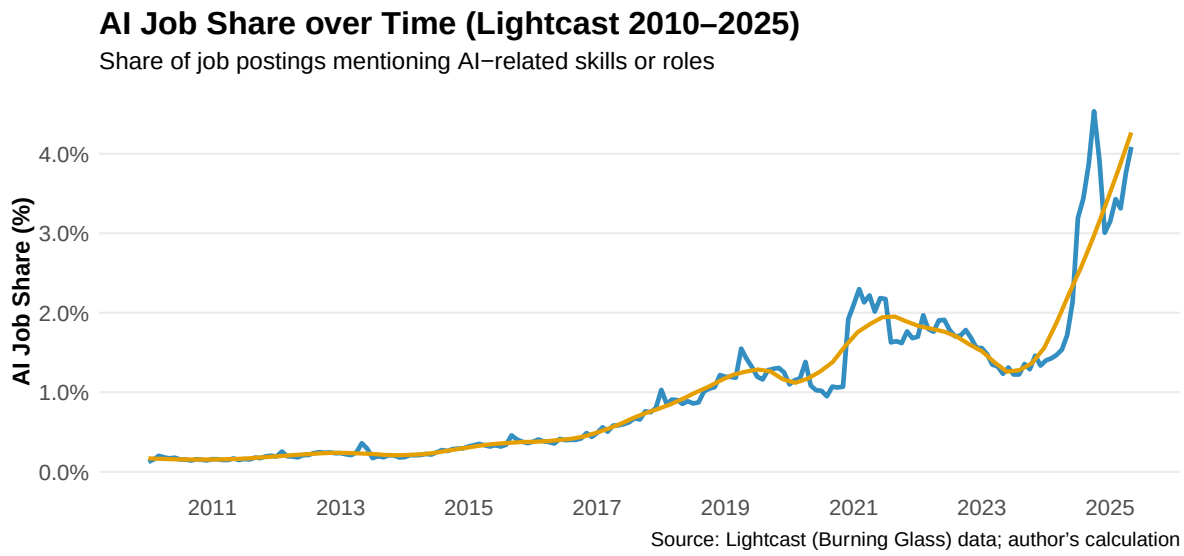

Figure A.2. Vertical composition

This figure plots the number of companies having each of the vertical listed. Note that a company can have multiple verticals; therefore, the histograms are not mutually exclusive.

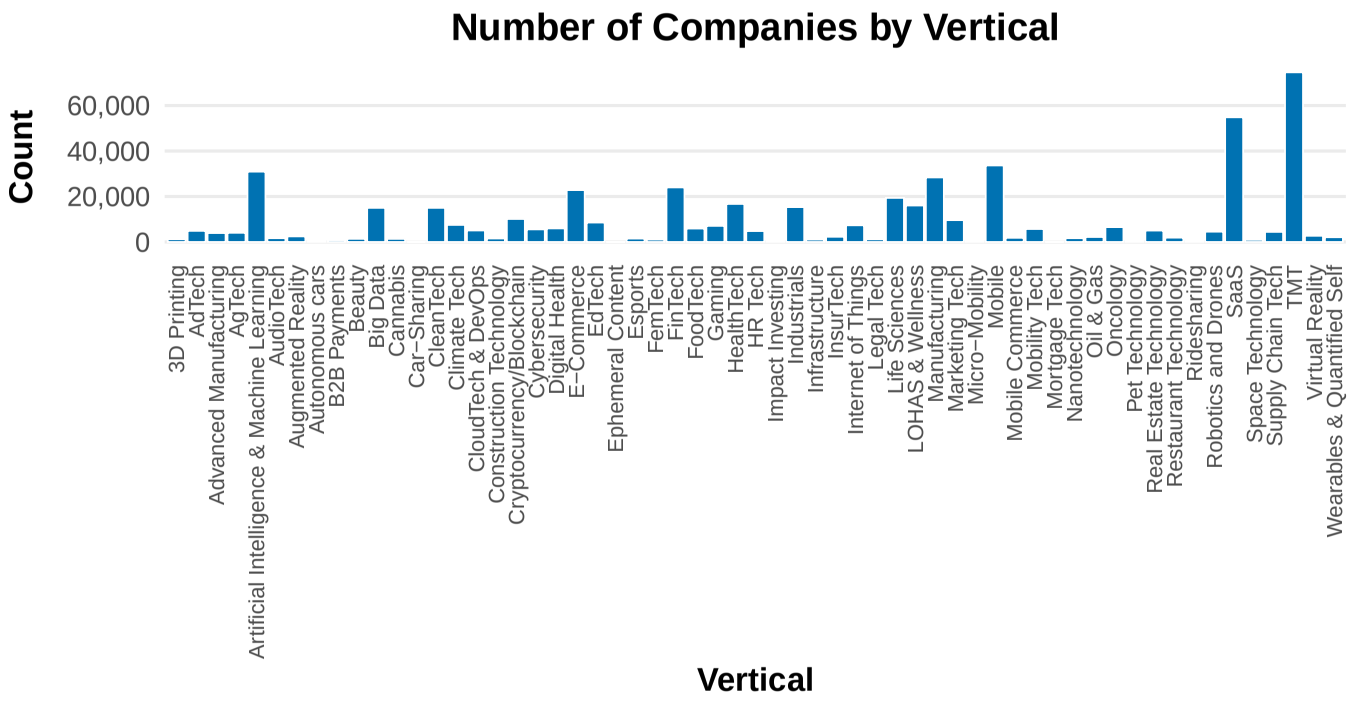

Figure A.3. Capital Share in AI (Primary Definition)

This figure shows the share of venture capital allocated to AI using PitchBook's primary vertical definition ("Artificial Intelligence & Machine Learning"). The top panel plots the money share, defined as the sum of deal sizes for AI startups divided by total VC deal value in each year. The bottom panel plots the deal share, defined as the number of AI deals divided by all VC deals in each year. Both measures show a substantial and sustained increase from 2010 to 2024, illustrating the growing concentration of venture activity in AI.

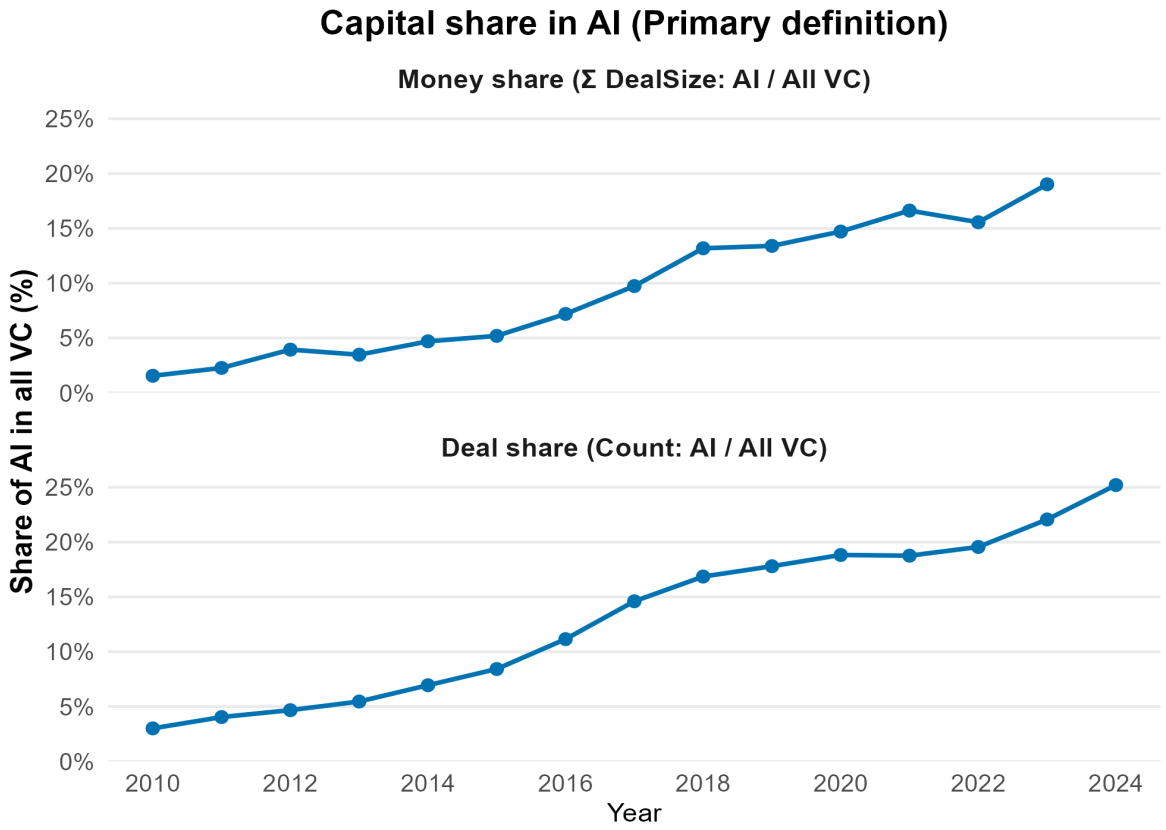

Figure A.4. Capital Share in AI (Primary Definition)

This figure shows the share of venture capital allocated to AI using PitchBook's primary vertical definition ("Artificial Intelligence & Machine Learning"). The top panel plots the money share, defined as the sum of deal sizes for AI startups divided by total VC deal value in each year. The bottom panel plots the deal share, defined as the number of AI deals divided by all VC deals in each year. Both measures show a substantial and sustained increase from 2010 to 2024, illustrating the growing concentration of venture activity in AI.

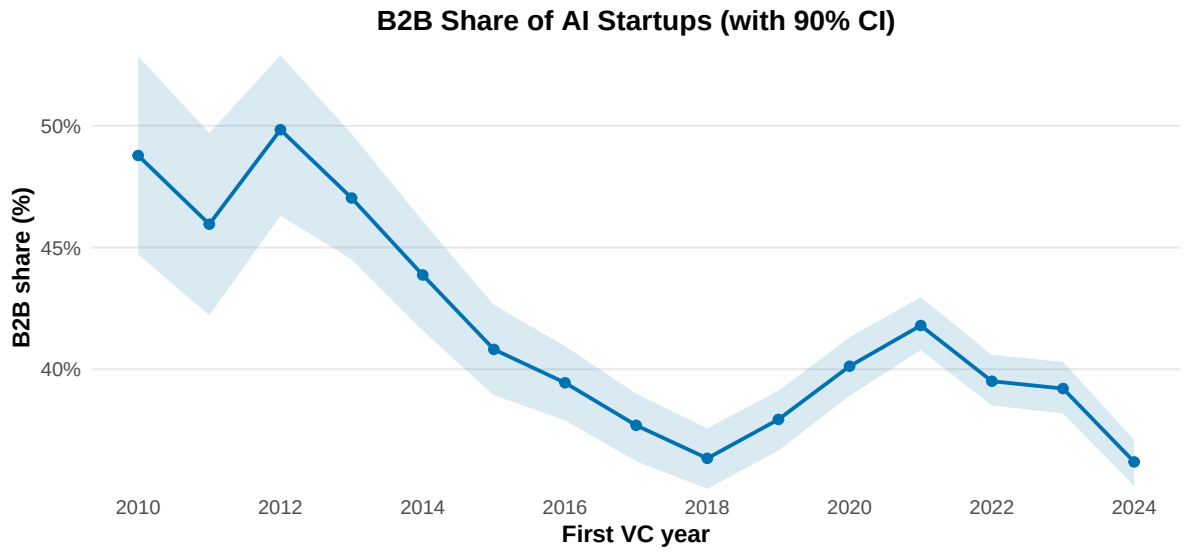

Figure A.5. Success rate by first VC year

This figure reports the cohort-level success rate of AI startups, defined as the fraction of firms achieving a successful exit (IPO or M&A with disclosed value exceeding total VC invested) based on their first venture financing year. Success rates are computed separately for B2B AI startups and Other AI startups using the strict exit definition. Across cohorts from 2010 to 2024, both groups exhibit a declining trend in exit likelihood, with B2B AI firms consistently showing slightly lower success rates, particularly among earlier cohorts.

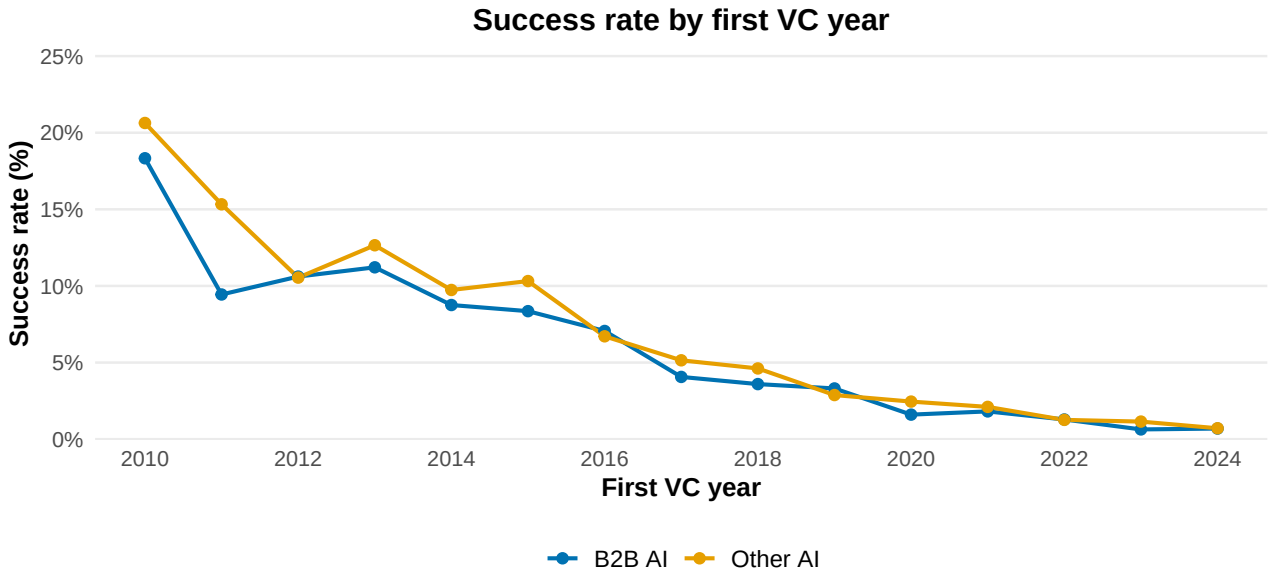

Figure A.6. Cause-specific hazard ratios for B2B vs. Other AI startups

This figure reports hazard ratios from cause-specific Cox proportional hazards models comparing the likelihood of exit for B2B AI startups relative to Other AI startups. The blue line shows the hazard ratio for IPO exits, and the yellow line shows the hazard ratio for M&A exits. Hazard ratios below 1 indicate a lower exit likelihood for B2B firms, while hazard ratios above 1 indicate a higher likelihood. B2B AI startups exhibit significantly lower IPO hazards and significantly higher M&A hazards, suggesting that B2B-oriented firms are more likely to be acquired rather than go public.

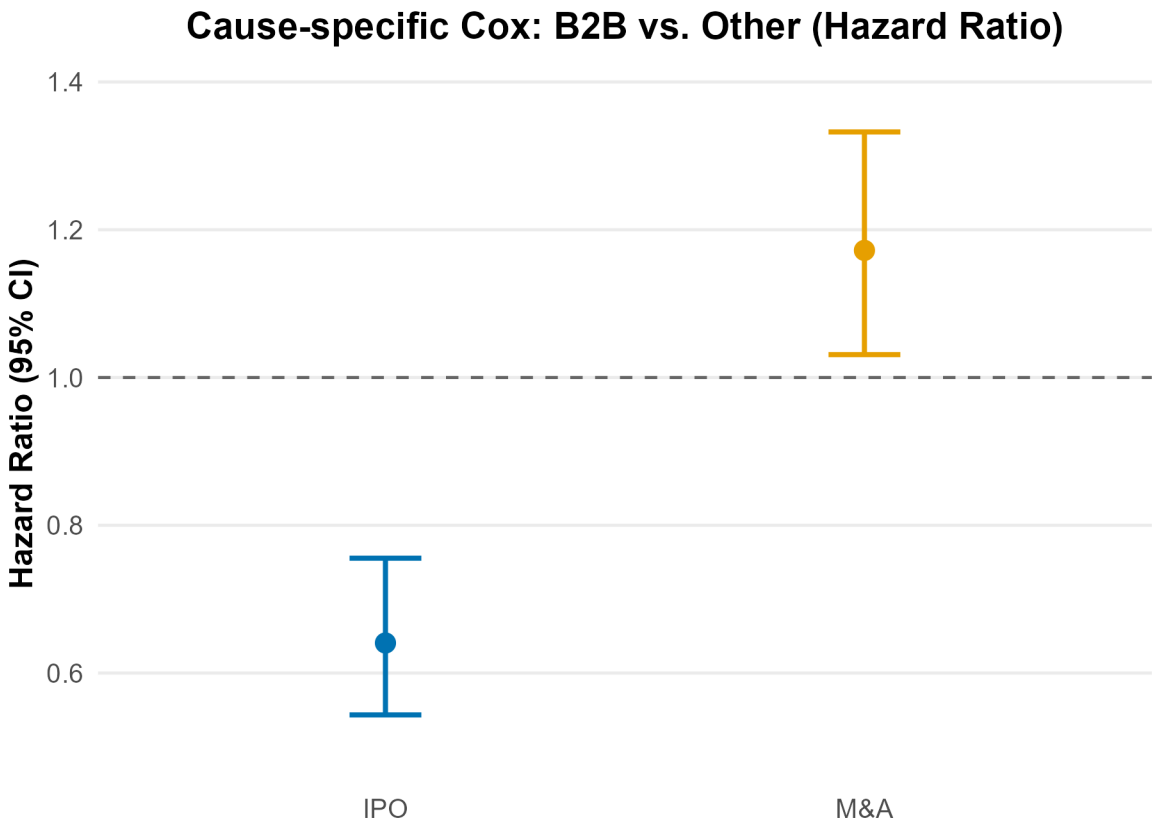

Figure A.7. Yearly average consistency of regulations by industry category

This figure illustrates the yearly average consistency trend across Federal Register documents including machine learning, human genome, global warming, fintech, cyber security, crypto, brain mapping, blockchain, and artificial intelligence as stem words between 2001 and 2024. Each line represents the average similarity score for a given category using either cosine similarity or Large Language Model (LLM), indicating the consistency or variability in regulatory focus within that category over time. Higher similarity scores reflect greater consistency in regulatory language and focus from year to year.

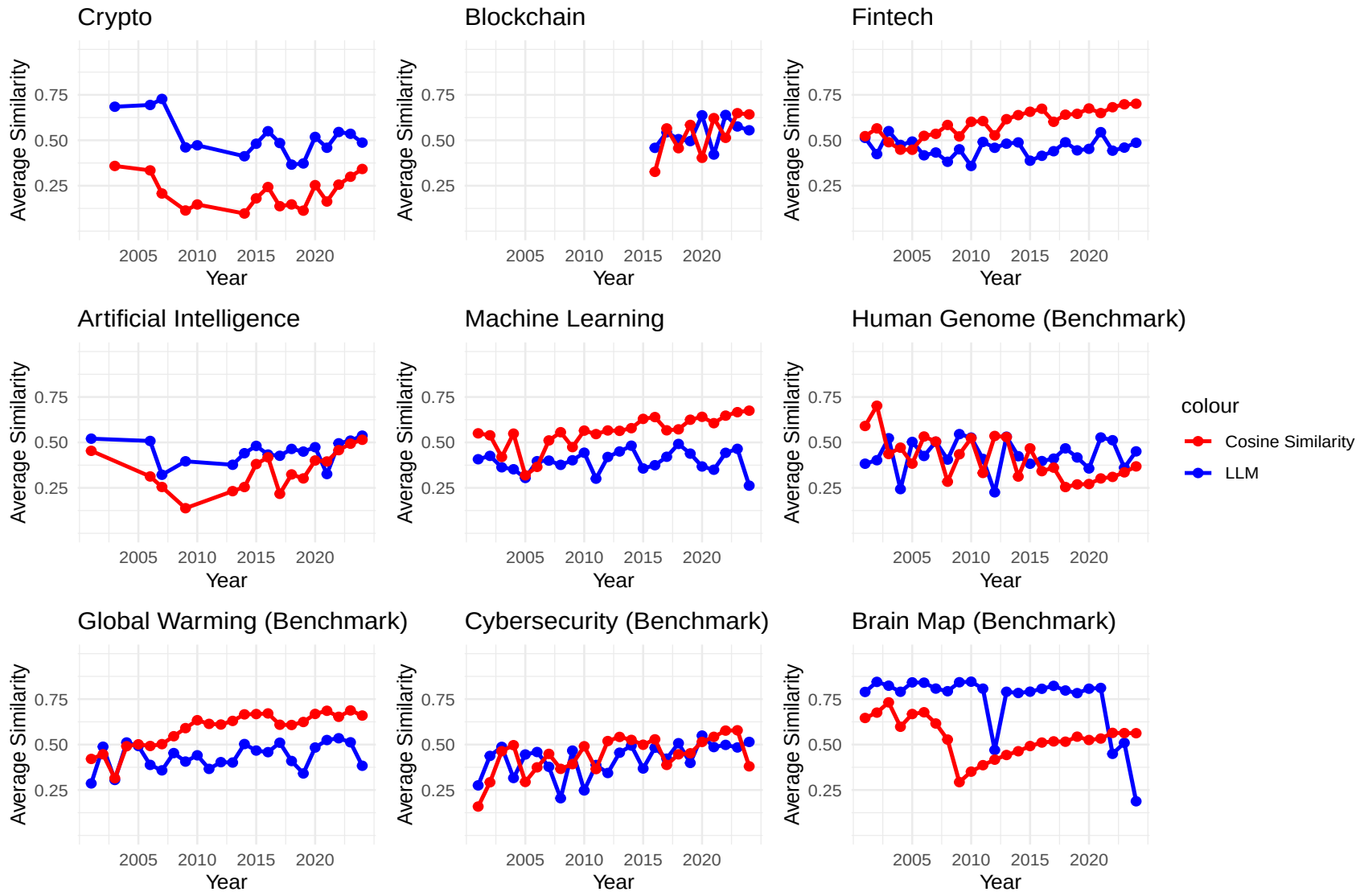

Figure A.8. Year wise consistency of regulations by industry category

This figure plots the year-to-year consistency trend across Federal Register documents including machine learning, human genome, global warming, fintech, cyber security, crypto, brain map, blockchain, and artificial intelligence as stem words between 2001 and 2024. Each cell represents a similarity score between documents from the years on the x- and y-axes using cosine similarity, capturing the consistency or shift in regulatory focus across time. Higher similarity scores, represented by darker colors, indicate greater consistency in regulatory language and themes within that category over time. Benchmark industry categories have a parentheses with B beside their names.

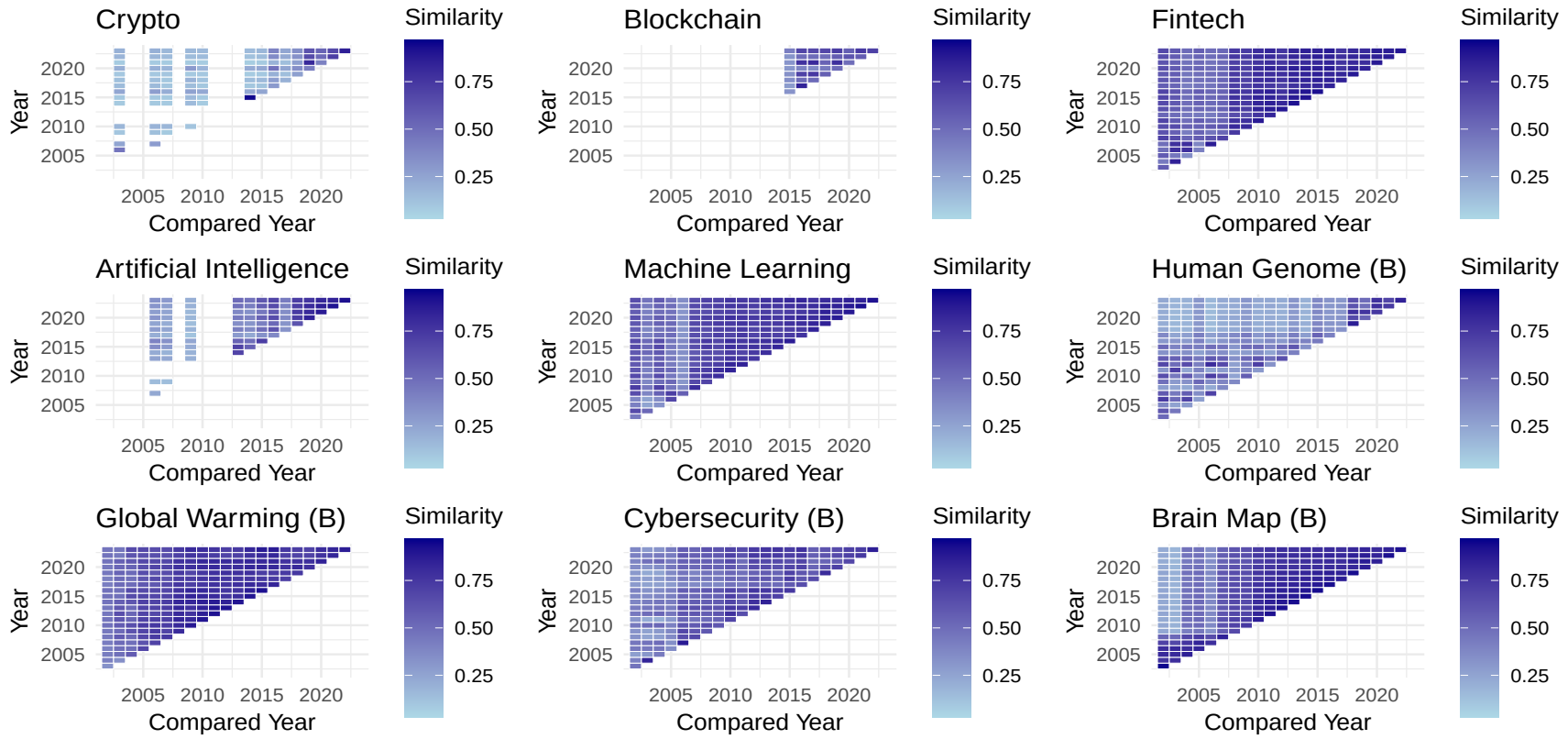

Table A.1. Top 15 AI-Related Job Titles

This table presents the 15 most common job titles associated with AI-related job postings matched to Compustat firms using Burning Glass Technologies data. The classification of AI jobs is based on occupation group, career area, ONET code, NAICS industry, and keyword matches in job titles.

Job Title	Count
Data Scientist	445
Software Engineer	412
Machine Learning Engineer	377
Data Analyst	365
Research Scientist	348
Ai Engineer	334
Senior Data Scientist	302
Computer Vision Engineer	294
Artificial Intelligence Lead	289
Nlp Engineer	275
Applied Scientist	264
Ml Engineer	258
Deep Learning Researcher	252
Ai Research Scientist	247
Data Science Manager	241

Table A.2. Security-Sensitive Healthcare Firms and Internal AI Hiring

This table presents regression results for a subsample of healthcare firms (SIC 8000–8099). The dependent variables include the AI job share, log of AI jobs, and a placebo test using non-AI job share. The key independent variable is *Security-sensitive (Health)*, an indicator for hospitals, nursing facilities, or health insurers (SIC codes 8062, 8051, and 6324). All regressions control for firm size (log(at)), ROA (sales/assets), leverage, and Tobin's Q. Standard errors are clustered by year.

	(1) AI Job Share	(2) Log(AI Jobs)	(3) Non-AI Job Share (Placebo)
Security-sensitive (Health)	0.008	0.586**	-0.008
	(0.008)	(0.156)	(0.008)
Log(at)	-0.004*	0.771***	0.004*
	(0.002)	(0.064)	(0.002)
ROA	-0.012	-2.651**	0.012
	(0.011)	(0.743)	(0.011)
Leverage	-0.010	-1.238*	0.010
	(0.007)	(0.480)	(0.007)
Tobin's Q	-0.002	0.079*	0.002
	(0.002)	(0.035)	(0.002)
Fixed Effects			
Year FE	Yes	Yes	Yes
Observations	777	319	777
R^2	0.066	0.485	0.066
R^2 (within)	0.016	0.446	0.016
RMSE	0.10	1.60	0.10

Standard errors clustered by year (fyear); t-stats in parentheses
*Significance Codes: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$*

A.3. Calculation of consistency scores (Federal Register documents) using cosine similarity and LLM

Document Classification. We start by removing false positives. Each document will be scored based on its relevance to its respective STEM word using the following criteria:
1. Score: 1
• Contains terms or topics directly or indirectly related to the STEM word.
• Examples include: “crypto asset, ” “digital asset, ” “virtual currency, ” “blockchain technology, ” or mentions of regulatory and security concerns specific to the cryptocurrency sector.
2. Score: 0.5
• Uses broader financial technology terms such as “digital finance” or “virtual assets” that imply cryptocurrency or blockchain relevance but lack direct specificity. (This is just an example based on cryptocurrency, but the method and idea remain consistent for other STEM terms.)
3. Score: 0
• Focuses solely on unrelated cryptographic technology, general encryption, or noncrypto contexts. (This is just an example based on cryptocurrency, but the method and idea remain consistent for other STEM terms.)

Code Implementation. The code classifies text into three categories based on the counts of direct terms and indirect terms related to cryptocurrency and blockchain. It assigns a score and a corresponding summary according to the following rules:
1. Condition 1 (Score: 1):
• If direct count $>$ 3 OR indirect count $>$ 9.
• Summary: *”Contains significant direct mentions of the given term.”*
2. Condition 2 (Score: 0.5):
• If 2 $<$ indirect count $<$ 10.
• Summary: *”Contains significant broader terms of the given term.”*
3. Condition 3 (Default, Score: 0):
• If neither of the above conditions is met.
• Summary: *”No relevant mentions or not significant enough of the given term.”*
Definitions of STEM terms and the corresponding classification terms (direct and indirect Words) are generated by ChatGPT.

TF-IDF with Cosine Similarity.
• Represents text using sparse vectors based on term frequency.
• Calculates similarity by measuring the angle between vectors.
• *Limitations:* Lacks semantic understanding; struggles with synonyms or word order.
• Example: “AI applications are increasing” and “Artificial intelligence usage is growing” might have low similarity.

SentenceTransformer-based LLM Model:
• Encodes text into dense vectors capturing semantic meaning.
• Uses transformer models (e.g., BERT) to understand word context.
• *Advantages:* Handles synonyms and word order well.
• Example: “AI applications are increasing” and “Artificial intelligence usage is growing” are similar due to semantic equivalence.

Table A.3. Consistency Scores Summary by Industry Category

This table provides a summary of the year-to-year similarity scores of regulatory language across different industry categories from Federal Register documents, including machine learning, human genome, global warming, fintech, cyber security, crypto, brain mapping, blockchain, and artificial intelligence between 2001 and 2024. The similarity scores capture the degree of consistency in regulatory focus within each industry over time. Panel A shows the statistics using a cosine similarity measure and Panel B shows the statistics using a LLM.

Panel A: Cosine Similarity

Industry	Mean	SD	Min	Q1	Median	Q3	Max
Crypto	0.2119	0.0879	0.0970	0.1444	0.1937	0.2668	0.3587
Blockchain	0.5292	0.1133	0.3263	0.4564	0.5653	0.6219	0.6491
Fintech	0.5938	0.0766	0.4479	0.5263	0.6038	0.6514	0.7012
Artificial Intelligence	0.3467	0.1092	0.1380	0.2547	0.3524	0.4269	0.5145
Machine Learning	0.5570	0.0891	0.3188	0.5441	0.5654	0.6260	0.6748
Human Genome	0.4105	0.1188	0.2544	0.3116	0.3757	0.5094	0.7018
Global Warming	0.5831	0.0976	0.3148	0.5020	0.6120	0.6613	0.6878
Cybersecurity	0.4432	0.1025	0.1589	0.3788	0.4577	0.5208	0.5784
Brain Map	0.5343	0.1078	0.2933	0.4847	0.5303	0.6025	0.7322

Panel B: LLM

Industry	Mean	SD	Min	Q1	Median	Q3	Max
Crypto	0.5157	0.1072	0.3660	0.4603	0.4859	0.5465	0.7274
Blockchain	0.5369	0.0747	0.4213	0.4954	0.5437	0.5759	0.6391
Fintech	0.4573	0.0474	0.3587	0.4304	0.4556	0.4884	0.5505
Artificial Intelligence	0.4475	0.0654	0.3220	0.4188	0.4569	0.4983	0.5373
Machine Learning	0.3955	0.0576	0.2622	0.3615	0.4008	0.4394	0.4917
Human Genome	0.4305	0.0848	0.2253	0.3936	0.4206	0.5056	0.5457
Global Warming	0.4345	0.0715	0.2862	0.3869	0.4470	0.4947	0.5347
Cybersecurity	0.4212	0.0902	0.2047	0.3753	0.4503	0.4865	0.5493
Brain Map	0.7434	0.1650	0.1879	0.7881	0.8022	0.8232	0.8460