When Firms Adopt AI Matters: Diffusion and the Value of AI Deployment

Parand Akbari[1]

Tepper School of Business, Carnegie Mellon University

First circulated: February 2026; Current version: July 2026.

Abstract

Does the value of adopting a general-purpose technology depend on how far it has diffused among competitors? A parsimonious real-options model formalizes two opposing forces: diffusion raises the competitive cost of remaining behind while compressing the scarcity value of firm-specific implementation capability. I identify the first strict public disclosure of AI deployment in 187,191 earnings and investor-call transcripts for 4,313 U.S. public firms from 2011 to 2024 and locate each adopter on a predetermined industry diffusion curve. Firms disclosing when peer deployment remains scarce revalue earlier: the first-mover-minus-follower difference in raw book-assets q is 0.45 in the first two post-disclosure years but narrows to 0.19 and becomes statistically indistinguishable from zero in years four and five. Continuous-diffusion and peer-catch-up tests yield the same ordering. The early differential survives capitalization of internally generated intangibles. Deployment shifts investment toward intangible capital, but operating returns do not show first-mover outperformance within the observed horizon. Competitive timing is therefore a first-order determinant of the market value associated with technology deployment.

Keywords: Artificial intelligence; technology diffusion; technology adoption; intangible capital; firm value; Tobin's q.

JEL: G14, G31, G32, O33, E22, L25.

[1] pakbari@andrew.cmu.edu

1 Introduction

The private value of a general-purpose technology depends on when a firm develops the complementary capability to use it relative to its competitors. Effective deployment can differentiate a firm while that capability remains scarce, but become necessary merely to preserve competitive parity once the technology has diffused. Electrification and information technology raised productivity only after firms reorganized production around them; the same organizational complements can make a technology differentiating early in its diffusion and a requirement for competitive parity later (David, 1990; Bresnahan et al., 2002; Bloom et al., 2012). Finance provides a forward-looking way to study that transition. If early deployment creates value because effective implementation remains scarce, the relative valuation advantage should be largest before competitors develop comparable capabilities and should narrow as the industry catches up.

Artificial intelligence is a sharp setting for this question because access is broad but effective deployment is not. Firms must combine models with proprietary data, software integration, technical talent, and redesigned workflows. These investments are firm-specific, costly, and largely intangible (Brynjolfsson et al., 2021, 2025). Thus two firms disclosing AI deployment in the same calendar year can face different economics. One may move while implementation capability is scarce; the other may adopt after AI has become a competitive requirement.

The empirical challenge is twofold: to identify concrete deployment rather than generic AI discussion, and to determine whether the valuation response reflects competitive timing rather than simply the market capitalization of complementary intangible investment. A higher q after deployment could reflect internally generated capital omitted from book assets, a generic growth option, or attention to the AI label. A competitive-timing mechanism makes a sharper prediction: the valuation response should be larger when peer deployment is low and should become less favorable as diffusion rises. I therefore use the industry's diffusion state to trace the component of the valuation response associated with competitive timing.

To organize these forces, I develop a parsimonious real-options model in which a

firm with pre-existing implementation capability chooses when to make an irreversible deployment and subsequently accumulates complementary intangible capital. Diffusion has opposing effects. It raises the competitive cost of remaining behind, which can induce defensive deployment by laggards, but it also reduces the differentiating payoff from deployment and the scarcity value of firm-specific capability. The model predicts earlier deployment by more capable firms, greater complementary intangible investment after deployment, and a valuation response that is strongest at low diffusion and erodes as peers catch up. These predictions organize the empirical analysis.

I construct a measure of the first strict public disclosure of AI deployment from the near-universe of S&P Capital IQ earnings and investor-call transcripts for U.S. public firms from 2011 to 2024. The final transcript universe contains 187,191 unique corporate events covering 4,313 firms. A large language model applies a fixed extraction schema pinned to the Capital IQ header firm, requires supporting quotations, and distinguishes concrete deployment from exploration and generic discussion. I align the exact disclosure date to Compustat fiscal periods and measure each adopter's position on a predetermined, lagged, leave-one-out industry diffusion curve. The empirical event is the first public disclosure of deployment, not the unobserved start of internal implementation. This design allows me to ask not simply whether valuation changes following deployment, but whether the response depends systematically on the competitive diffusion state at which the firm enters. To summarize this heterogeneity, I classify adopters at or below the within-industry 25th percentile of lagged leave-one-out peer diffusion as first movers and the remaining adopters as followers.

The central result is that competitive timing governs the valuation path following disclosed AI deployment. Over the first two post-disclosure years, the first-mover-minus-follower difference in raw book-assets q is $+0.447$. This difference reflects both an earlier positive revaluation among first movers, whose estimated effect is $+0.177$, and an initial decline among followers, whose estimated effect is -0.270 relative to their untreated counterfactual. By years four and five, both groups have positive estimated effects—$+0.434$ for first movers and $+0.242$ for followers—and the difference has narrowed to

+0.192, statistically indistinguishable from zero. Firms that disclose while deployment remains scarce therefore revalue earlier; the evidence does not indicate a permanently higher valuation level for first movers.

The early differential is not explained by the omission of internally generated capital from conventional book assets. It remains +0.367 when estimated knowledge and organization capital are added to the asset denominator and +1.165 on Peters–Taylor total q, with similar results across alternative capitalization parameters (Peters and Taylor, 2017). The same economic ordering appears when diffusion is measured continuously rather than through the first-mover split. Post-disclosure valuation responses are less favorable at higher predetermined diffusion on the two intangible-capitalized measures. Within first movers, a ten-percentage-point increase in subsequent peer diffusion is also associated with approximately 0.10 lower raw q over event years one through five. These results connect the valuation pattern to the competitive diffusion state rather than to a particular quartile cutoff or accounting perimeter.

Additional evidence maps the valuation result to the paper's conceptual framework. Consistent with the model, pre-disclosure capability is more strongly associated with valuation when diffusion is low in the baseline specification. After disclosure, investment shifts toward R&D and booked intangibles rather than physical capital, but first movers do not outperform followers on any outcome in the pre-specified operating family within the observed horizon. These findings are consistent with capability-intensive adjustment whose value is priced before superior operating performance becomes visible in the accounting data.

Taken together, the paper establishes a diffusion-sensitive valuation fact: disclosed AI deployment is valued more favorably when peer deployment remains scarce, and the relative advantage attenuates as competitors catch up. The evidence is consistent with a temporary scarcity value of firm-specific implementation capability. Whether that value is realized through operating rents or longer-duration growth opportunities cannot yet be distinguished, but competitive diffusion is a first-order state variable for valuing technology deployment.

This paper contributes to three literatures. First, the paper contributes a dated and auditable measure of operational technology deployment. Existing measures capture task exposure, patents, AI hiring, generic disclosure, and survey-reported use (Eloundou et al., 2023; Babina et al., 2023, 2024; McElheran et al., 2024). Concurrent and subsequent work further characterizes firm-level AI use through executive surveys, nationally representative business data, and the depth of integration reported in periodic filings (Yotzov et al., 2026; Bonney et al., 2026; Yu et al., 2026). The present measure instead identifies the first strict public disclosure of realized deployment, supplies the supporting text, and locates each firm on its industry diffusion curve at that date.

Second, the paper shows that the valuation associated with technology deployment depends on the competitive diffusion state. Existing finance research studies AI exposure, hiring-based investment, and market responses to broad technology shocks (Eisfeldt et al., 2023; Babina et al., 2024). By dating firm-specific deployment and locating each firm on its industry diffusion curve, I show that deployment is revalued earlier when effective implementation remains scarce and that the relative differential narrows as peers catch up. The result survives capitalization of internally generated intangibles, connecting technology diffusion to the finance of organization capital and total q (Eisfeldt and Papanikolaou, 2013, 2014; Peters and Taylor, 2017; Corrado et al., 2021).

Third, the paper provides firm-level evidence on strategic technology timing. Theoretical work on first-mover advantage, preemption, and strategic option exercise emphasizes that the value of investment depends on rivals' actions (Lieberman and Montgomery, 1988; Fudenberg and Tirole, 1985; Grenadier, 1996). The empirical contribution here is to measure that competitive state directly: the same broad technology is associated with different valuation paths depending on how far peer deployment has progressed.

Section 2 develops and validates the deployment measure. Section 3 presents the conceptual framework. Section 4 defines the fiscal alignment, valuation perimeters, and empirical designs. Section 5 documents the diffusion-dependent valuation response. Section 6 examines capitalization, capability, and realized operating returns. Section 7 reports complementary investment and balance-sheet adjustment. Section 8 concludes.

2 Measuring Disclosed AI Deployment

2.1 Transcript universe

I assemble a panel of English, earnings-related transcripts from S&P Capital IQ Transcripts spanning 2011–2024. Restricting the source file to events dated 2011–2024 and retaining one observation per corporate-event identifier yields 187,191 unique corporate transcript events for U.S. public companies. [2] Each transcript is parsed at the firm–event level using Capital IQ's header Target Company identifier. When a transcript discusses multiple firms (for example, suppliers or peers), I extract information only for the header firm by design. Each observation corresponds to one Capital IQ transcript event for the header firm, timestamped at the event date and time in the vendor metadata. Panel A of Figure 1 summarizes the annual evolution of the transcript universe and the AI-related subsets used in the analysis. Internet Appendix Tables IA.1 and IA.2 report the corresponding annual counts and event-type composition.

2.2 From transcripts to a deployment score

I convert transcripts into structured firm–event observations in two steps. A deterministic keyword screen routes potentially AI-relevant transcripts to a large language model (LLM), which returns an evidence-gated ordinal score for general-AI deployment, $AID_{i,e} \in \{0, 0.5, 1\}$, for transcript event e of firm i. A score of 1 denotes concrete deployment or integration, 0.5 denotes exploratory, pilot, or aspirational activity, and 0 denotes no deployment content. The baseline treatment date is the first transcript event with $AID = 1$. I interpret this date as the first public evidence of deployment, which may follow internal experimentation or implementation. The pipeline also constructs an analogous generative-AI score, $GenAID_{i,e}$, for descriptive and validation exercises; the baseline treatment throughout the paper is defined using the general-AI score $AID_{i,e}$.

The deterministic keyword screen produces 23,459 raw transcript–company records. Because the source contains alternate transcript versions for some corporate events, I

[2] Private and non-U.S. companies are excluded, so the panel should be read as representative of U.S. public issuers covered by Capital IQ.

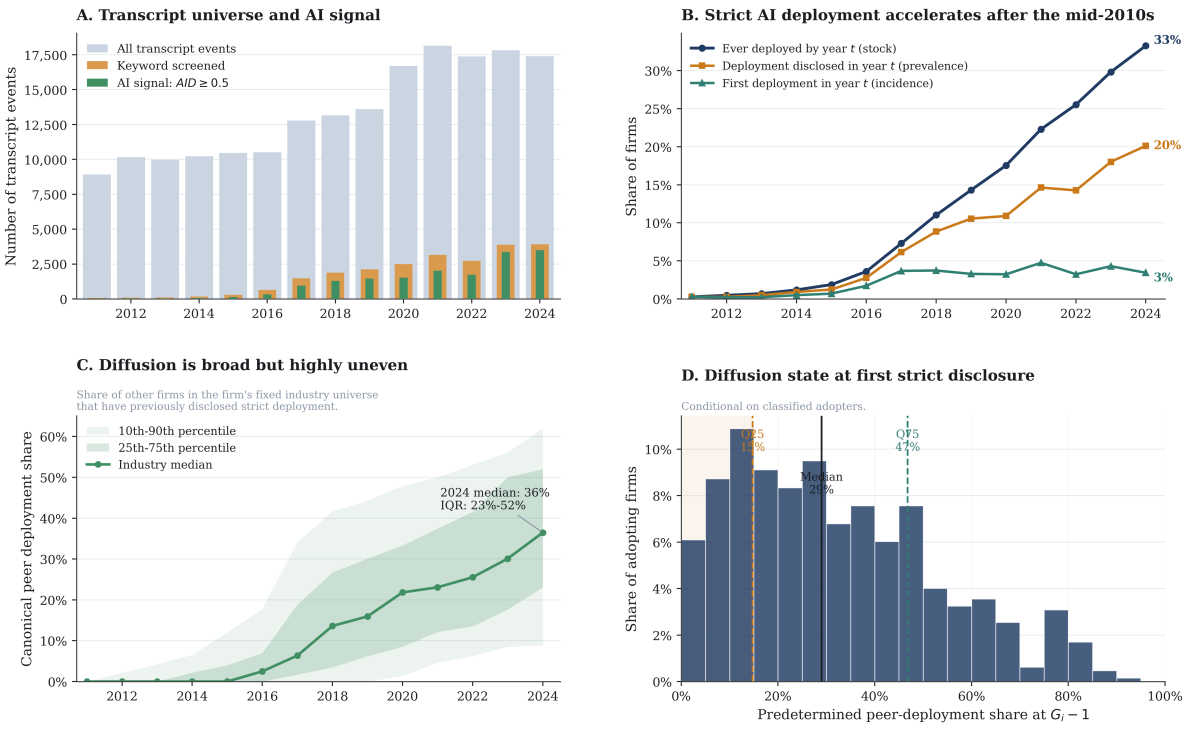


Figure 1: **Disclosed AI deployment and diffusion.** Panel A reports annual counts of unique corporate transcript events, the deduplicated keyword-screened subset, and deduplicated events containing an AI deployment or exploration signal ($AID \geq 0.5$). Panel B reports strict deployment ($AID = 1$) in the full transcript-firm universe: the cumulative share of firms that have ever disclosed deployment by year t, the share with a strict deployment disclosure during year t, and the share first disclosing strict deployment during year t. Panel C reports, for each calendar year, the equally weighted cross-industry distribution of peer deployment. The denominator for each industry is the other firms in its fixed industry universe; the figure shows the industry median, interquartile range, and 10th–90th percentile range. Panel D conditions on the 1,293 stock-aligned adopters and reports the distribution of the predetermined leave-one-out peer-disclosure state measured immediately before each firm's first strict disclosure. The vertical lines mark the pooled 25th percentile, median, and 75th percentile (15%, 29%, and 47%). These are diffusion states faced by classified adopters, not aggregate adoption rates.

collapse records sharing the same corporate-event identifier before constructing the event-level series. This leaves 22,974 unique screened corporate events, all of which carry an AID score. Of these events, 16,376 have $AID \geq 0.5$ and 12,617 have $AID = 1$. Table A.1 summarizes the aggregate measurement pipeline and the resulting analysis samples.

Evidence and prompt design. I run one prompt per (transcript, company) pair. The prompt constrains every answer to the Capital IQ header target company and returns a JSON object that populates the analysis variables. Every nonzero AID or GenAID score must be accompanied by one or more supporting sentences from the transcript.

Other extracted fields. The prompt also records (i) role and sourcing (Developer, Adopter, or Both; In-house, External, or Both; and vendor names when disclosed), (ii) use case and scope (core versus supportive, application types, and business units), and (iii) intensity and depth (integration level, investment level, extent of adoption, and model names when mentioned). Detailed prompt design, the scoring protocol, and representative examples are provided in Internet Appendix Section IA.1.3.

2.3 Reliability and validation

Transcript components are streamed and concatenated, identifiers are normalized, and rate limits, checkpoints, and retries ensure stable processing. The model is constrained to return JSON only, and non-parseable responses are retried with a minimal wrapper. To guard against hallucination, the prompt includes three placebo questions (a fictional AI use case, a secret AI department, and a quantum-AI investment) that should return "Information not found" when the transcript does not support them. I also manually audit a random subset of events, verifying both that the score is correct and that the supporting quote refers to the header firm rather than to a third party mentioned in the transcript. Extractions were run with Claude 4 Sonnet in JSON mode between July and August 2025.

Table 1 summarizes the validation diagnostics. Agreement with human coding is high: three-class accuracy is 0.914, and the strict adopter class ($AID = 1$) has precision 0.993.

Supporting quotes anchor to the transcript in 97.5% of cases under the strict definition and 99.1% under the broader locatability definition, while abstention exceeds 99.5% for the two strict placebos. Internet Appendix Section IA.1.4 reports the underlying anchoring, placebo, and confusion-matrix results.

Table 1: LLM adoption measure: validation summary

Diagnostic	Value
Manual audit vs. human coding ($N = 455$)	
AI adoption, three-class accuracy	0.914
Strict adopters ($AID = 1$): precision / recall	0.993/0.924
Binary AI adoption ($AID \geq 0.5$): precision / recall	0.989/0.962
Binary GenAI adoption ($GenAID \geq 0.5$): precision / recall	0.908/0.975
Evidence auditability (raw parsed-record sample, $N = 23{,}459$)	
Supporting-quote anchoring (strict / locatable)	97.5%/99.1%
Placebo abstention (two strict placebos)	$\geq 99.5\%$

Notes: The human label is the reference. The intermediate 0.5 category is intentionally noisier and is not the baseline treatment (the baseline uses the strict threshold $AID = 1$). Underlying diagnostics are reported in Internet Appendix Section IA.1.4.

2.4 Linkage, coverage, and the observation process

Firm–event extractions are merged to standard firm identifiers (Capital IQ companyid to gvkey when available) and then to Compustat outcomes and GICS industry definitions. I construct annual firm panels spanning 2011–2024.

Aggregation to firm–year measures. I aggregate the event-level extractions into analysis-ready firm–year variables. These include AI and GenAI disclosure indicators and first-disclosure timing, role and sourcing tags, application and business-unit tags, and measures of integration intensity. Binary indicators are aggregated using a logical-any rule, while graded measures are aggregated using means or sums, as appropriate. Compustat variable definitions and fill rules are reported in Appendix A, and detailed construction of the nonstandard accounting outcomes is reported in Internet Appendix Section IA.2.

Coverage and the observation process. Within the estimation panel used in the event studies, transcript coverage ranges from roughly 68% to 92% across years and rises to

92% by 2024. Pooled over the broader coverage panel of all 4,313 firms across 2011–2024, the share of firm–years containing a transcript is mechanically lower, at 55%, because the balanced grid includes inactive and pre-entry cells. The corresponding coverage distribution is reported in Internet Appendix Figure IA.2.

The date of the first strict disclosure is observed exactly, but the underlying transition to deployment is interval censored: it occurs sometime between the firm's last transcript with $AID < 1$ and its first transcript with $AID = 1$.[3] The median length of this interval is 160 days. A firm–year without a transcript cannot reveal a new disclosure. Accordingly, never adopters are firms that never disclose strict deployment through 2024, not firms verified never to have deployed AI. The baseline analysis uses the full estimation panel and treats the first observed strict disclosure as the event date; its alignment to fiscal periods is defined in Section 4.1.

To assess whether the results are sensitive to the transcript observation process, I construct three alternative samples. The first two modify the set of untreated observations. The current-period-observable sample retains an untreated firm–period only when the firm has a transcript in that period. The right-censored sample is less restrictive: it permits temporary gaps in transcript coverage but excludes untreated firm–periods occurring after the firm's final observed transcript. These restrictions test whether the results depend on retaining periods in which the data provide no credible opportunity to observe a new disclosure.

The third sample addresses treatment-date precision rather than the untreated comparison group. It retains all never-disclosing firms but keeps an adopter only when the interval between its last observed non-strict transcript and its first strict disclosure is nonmissing and no more than 365 days. This restriction removes adopters whose transition to deployment is only loosely dated.

Figure A.1 summarizes transcript coverage, disclosure intervals, and event-horizon retention, where retention is the share of calendar-eligible adopters with an observed outcome at each event horizon. Tables A.2 and A.3 report the corresponding sample

[3]This interval need not bracket the start of internal implementation.

definitions and retention rates. The estimates under the three alternative samples are reported in Internet Appendix Section IA.6.

2.5 Who adopts AI and who moves first

AI deployment is systematically selected, although observable characteristics predict subsequent deployment only moderately out of sample. I estimate a discrete-time hazard model using fiscal firm–years in which the firm has at least one transcript. For firms that eventually disclose, the dependent variable equals one in the first fiscal year whose year-end falls on or after the first strict disclosure and zero in earlier observed years. For never-disclosing firms, it equals zero throughout. Observations after a firm's first-disclosure year are excluded, and firm characteristics enter with a one-year lag. The strongest correlate is the lagged leave-one-out industry deployment share: a one-standard-deviation increase is associated with 2.40 times the odds of first strict disclosure (95% CI $[1.89, 3.04]$). Cash/assets, booked intangibles/assets, and firm size are also positively associated with first disclosure, with odds ratios of 1.27, 1.16, and 1.13, respectively. The remaining accounting predictors are not statistically distinguishable from zero. Internet Appendix Table IA.8 reports the complete hazard specification.

A ridge-logit model containing only calendar-year and industry effects achieves a firm-grouped cross-validated AUC of 0.626. Adding lagged fundamentals and the diffusion state raises the AUC to 0.656, while a fixed calendar holdout that trains through 2020 and evaluates 2021–2024 produces a similar AUC of 0.653. Thus deployment responds strongly to the industry diffusion environment, but observable firm fundamentals add only modest prospective predictive content. The prediction design and out-of-sample validation procedure are detailed in Internet Appendix Section IA.3.

Selection also operates within the set of eventual adopters. Under the predetermined first-mover/follower split introduced above and formalized in Section 4.4.1, first movers reach their first strict deployment disclosure with a larger accumulated stock of intangible capital. Their pre-adoption log constructed intangible-capital stock is 0.27 standard deviations higher than that of followers after controlling for six-digit industry effects, and

remains 0.21 standard deviations higher after additionally controlling for pre-adoption firm size. They also have higher capital-expenditure intensity in the size-adjusted specification. However, first movers do not have significantly higher R&D/assets, booked intangibles/assets, or constructed intangible capital/assets. Thus, first movers enter disclosure with a larger accumulated stock of constructed intangible capital, even after controlling for firm size, but not with significantly higher R&D intensity or asset-scaled intangible holdings. Internet Appendix Figure IA.5 reports the full set of pre-adoption comparisons.

These comparisons are descriptive and do not make adoption or diffusion position randomly assigned. They motivate the paper's reliance on within-firm dynamics and on the predetermined continuous-diffusion gradient rather than on treated-versus-untreated level comparisons.

2.6 Adoption dynamics

The deployment measure produces broad but uneven diffusion across firms and industries. Panels B–D of Figure 1 summarize these patterns using strict deployment ($AID = 1$) aggregated to the firm–year level.

Panel B shows that strict deployment accelerates after the mid-2010s. The share of firms that have ever disclosed deployment rises from near zero before 2016 to 33% by 2024. The share disclosing deployment in a given year reaches 20%, while the share disclosing for the first time peaks near 4% and settles around 3%. Thus, the cumulative stock of disclosers grows steadily even though the annual flow of new disclosers remains modest.

Panel C shows that diffusion is broad but highly uneven across industries. By 2024, the median industry peer-deployment share is 36%, with an interquartile range from 23% to 52%.

Panel D changes the unit of observation from the industry–year to the individual adopter. For each adopter, I calculate the share of other firms in the same industry that had already disclosed strict deployment by the end of the year preceding the focal firm's first strict disclosure. This lagged, leave-one-out peer share is the predetermined diffusion state used in the paper's heterogeneity designs. The diffusion state varies substantially

across adopters. At the median adopter's first strict disclosure, 29% of its industry peers had already disclosed strict deployment by the end of the preceding year. One quarter of adopters disclose when the corresponding peer share is below approximately 15%, while one quarter disclose when it exceeds approximately 47%.

Panel C therefore describes the evolution of industry diffusion in calendar time, whereas Panel D locates each adopter on its industry's diffusion curve at the time of its own disclosure. The cross-sectional variation in Panel D is the variation used by the first-mover split (Section 4.4.1) and the continuous-diffusion design (Section 4.4.3). Further evidence on implementation depth, industry composition, and the distinct GenAI diffusion wave is reported in Internet Appendix Section IA.4.

3 A Conceptual Framework

This section develops a stylized partial-equilibrium model to organize the empirical analysis. The model is not structurally estimated; its role is to clarify the economic mechanisms and generate comparative predictions examined in the data. An observed valuation response to deployment can combine several distinct components: the option to delay deployment, the value of complementary intangible capital installed after deployment, future growth and adjustment opportunities, and the temporary scarcity value of deploying before effective implementation has diffused widely. The framework also shows why diffusion can encourage defensive deployment by firms that remain behind while reducing the value associated with deployment conditional on having deployed.

3.1 Deployment opportunities, diffusion, and operating payoffs

Time is continuous. A firm has predetermined implementation capability $a > 0$, summarizing the data, software, technical personnel, and organizational readiness in place before deployment. Industry diffusion is $S_t \in [0, 1]$ and evolves according to

$$\dot{S}_t = m(S_t) \geq 0, \qquad m(1) = 0. \tag{1}$$

Diffusion is therefore weakly increasing, mirroring the absorbing peer-disclosure share used in the empirical analysis. The firm is small relative to its industry and takes the diffusion path S_t as given.

The profitability of the AI application currently available to the firm is summarized by a firm-specific state $X_t > 0$ satisfying

$$dX_t = \mu_X X_t \, dt + \sigma_X X_t \, dW_t, \qquad \sigma_X > 0. \tag{2}$$

A high value of X_t represents an application with a strong technical fit, large expected cost savings, or substantial revenue potential; a low value represents an opportunity that is not yet sufficiently profitable to justify deployment. Uncertainty about X_t creates an option to wait for a more favorable application opportunity. The process is specified under a reduced-form risk-adjusted valuation measure, and r denotes the corresponding constant discount rate.

Before deployment, flow profit is

$$\pi^N(a, S) = \bar{\pi}(a) - L(S), \qquad L'(S) > 0, \tag{3}$$

where $L(S)$ is the competitive loss from remaining behind as deployment becomes common among industry peers.

Deployment is irreversible and requires a sunk implementation cost $F(a)$, with $F'(a) < 0$. Greater capability therefore makes implementation less costly. Deployment also installs an initial stock of complementary intangible capital $K_0(a) > 0$, with $K_0'(a) \geq 0$, and changes flow profit to

$$\pi^A(a, K, S, X) = \bar{\pi}(a) + aXg(S)f(K), \tag{4}$$

where

$$f' > 0, \qquad f'' < 0, \qquad g > 0, \qquad g' < 0.$$

The term $aXg(S)f(K)$ is the differentiating payoff from deployment. Implementation

capability a, a more profitable application opportunity X, and a larger stock of complementary capital K raise that payoff. Diffusion reduces its differentiating component because effective AI deployment becomes closer to an industry-parity input. The restriction $g'(S) < 0$ should be read as the net effect of diffusion on the private differentiating payoff, after any diffusion-related improvements in common infrastructure or vendor availability.

3.2 Post-deployment complementary investment

After deployment, the firm chooses gross investment $I_t \geq 0$ in complementary intangible capital:

$$\dot{K}_t = I_t - \delta K_t, \qquad C(I_t, K_t) = I_t + \frac{\gamma}{2}\frac{I_t^2}{K_t}, \qquad \delta, \gamma > 0. \tag{5}$$

The capital stock represents assets such as data architecture, software integration, technical systems, and organizational capital. The restriction $I_t \geq 0$ captures costly irreversibility: installed complementary capital cannot be costlessly sold back or dismantled.

Conditional on having deployed, firm value is

$$V^A(a, K, S, X) = \sup_{\{I_t\}_{t\geq 0} \in \mathcal{A}(a,K,S,X)} \mathbb{E}_{K,S,X}\left[\int_0^\infty e^{-rt}\left\{\pi^A(a, K_t, S_t, X_t) - C(I_t, K_t)\right\} dt\right], \tag{6}$$

where $\mathcal{A}(a, K, S, X)$ is the set of admissible nonnegative investment policies. The technical conditions ensuring that the state processes and value function are well defined are stated in Internet Appendix Section IA.5.

For a sufficiently smooth function h, define the generator of the uncontrolled states by

$$\mathcal{L}h = m(S)h_S + \mu_X X h_X + \frac{1}{2}\sigma_X^2 X^2 h_{XX}.$$

The dynamic-programming equation for post-deployment value is

$$rV^A = \pi^A + \max_{I\geq 0}\left\{-I - \frac{\gamma}{2}\frac{I^2}{K} + V_K^A(I - \delta K)\right\} + \mathcal{L}V^A. \tag{7}$$

Let

$$q^K \equiv V_K^A$$

denote the shadow value of one additional unit of installed complementary capital. The optimal investment rule is

$$\frac{I^*}{K} = \max\left\{0, \frac{q^K - 1}{\gamma}\right\}. \tag{8}$$

The firm does not invest when the shadow value of another unit of capital is below its unit purchase cost. Once q^K exceeds one, investment intensity increases with marginal q^K. Thus deployment can initiate a period of complementary investment rather than produce its full operating payoff immediately.

For additional intuition, temporarily hold S and X fixed. At an interior stationary target, replacement investment satisfies $I = \delta K$, and the target stock $K^*(a, S, X)$ solves

$$aXg(S)f'(K^*(a, S, X)) = (r + \delta)(1 + \gamma\delta) - \frac{\gamma}{2}\delta^2 \equiv \chi > 0. \tag{9}$$

Because $f'' < 0$, the target stock is increasing in capability a and application profitability X, and decreasing in diffusion S. Along the actual path, however, both S_t and X_t change, so $K^*(a, S, X)$ is a moving target rather than a literal long-run steady state. Internet Appendix Section IA.5 provides the derivation and additional comparative statics.

3.3 The option to wait and the deployment boundary

The value of deploying immediately is

$$J(a, S, X) = V^A(a, K_0(a), S, X) - F(a). \tag{10}$$

Before deployment, the firm chooses a stopping time τ:

$$V^N(a, S, X) = \sup_{\tau} \mathbb{E}\left[\int_0^{\tau} e^{-rt}\pi^N(a, S_t)\, dt + e^{-r\tau}J(a, S_\tau, X_\tau)\right]. \tag{11}$$

Equivalently, pre-deployment value satisfies

$$\max\left\{\pi^N + \mathcal{L}V^N - rV^N,\ J - V^N\right\} = 0. \tag{12}$$

In the continuation region, the value of waiting exceeds the value of deploying immediately. In the stopping region, immediate deployment is optimal.

Internet Appendix Section IA.5 states sufficient monotonicity conditions under which the advantage of deploying rather than waiting increases with the application-profitability state X. Under these conditions, the optimal policy has a threshold form:

$$X \geq X^*(a, S). \tag{13}$$

The firm deploys when the profitability of its available application reaches the threshold $X^*(a, S)$ and waits below that threshold. At a regular boundary, value matching and smooth pasting imply

$$V^N(a, S, X^*) = J(a, S, X^*), \tag{14}$$

$$V^N_X(a, S, X^*) = J_X(a, S, X^*). \tag{15}$$

Greater capability raises the payoff from deployment, lowers the sunk implementation cost, and weakly increases the initial stock of complementary capital. Under the maintained monotone-deployment condition, the deployment threshold $X^*(a, S)$ therefore weakly falls with a. Capability makes deployment more likely, but it does not make deployment automatic: even a highly capable firm can optimally wait when the currently available application is insufficiently profitable.

Diffusion affects deployment through two opposing channels. First, it reduces the differentiating payoff from AI through $g'(S) < 0$. Second, it raises the competitive loss suffered by firms that remain behind through $L'(S) > 0$. Because deployment allows the

firm to avoid that loss, the gain from deploying rather than remaining a non-adopter is

$$\Delta\pi(a, K, S, X) = L(S) + aXg(S)f(K), \qquad \Delta\pi_S = L'(S) + aXg'(S)f(K). \tag{16}$$

The first term in $\Delta\pi_S$ is an avoided-loss channel: as peer deployment rises, remaining a non-adopter becomes increasingly costly. The second is a differentiation channel: broader diffusion reduces the private payoff from possessing a capability that competitors do not yet have.

When the avoided competitive loss dominates in present value, diffusion lowers the deployment threshold and can induce laggards to deploy. At the same time, diffusion reduces the differentiating value of deployment for firms that have already deployed. Diffusion can therefore make deployment increasingly necessary while making it less valuable as a source of competitive advantage.

3.4 Temporary scarcity value and empirical predictions

Define the temporary scarcity value of incomplete diffusion as

$$\mathcal{R}(a, K, S, X) = V^A(a, K, S, X) - V^A(a, K, 1, X). \tag{17}$$

This object compares the value of the same deployed firm at the current diffusion state with its value when industry deployment is already saturated. Under the order-preserving and decreasing-differences conditions stated in Internet Appendix Section IA.5,

$$\mathcal{R} \geq 0, \qquad V^A_S \leq 0, \qquad V^A_{aS} \leq 0. \tag{18}$$

The first two inequalities state that incomplete diffusion can generate temporary scarcity value and that post-deployment value weakly declines as diffusion rises. The third states that diffusion also reduces the incremental value of implementation capability. Capability is therefore particularly valuable when effective deployment remains uncommon and less differentiating once competitors have caught up.

The sign of V_{aS}^{A} is conditional rather than automatic. Holding the complementary capital stock fixed, current flow profit has decreasing differences in capability and diffusion:

$$\frac{\partial^2 \pi^A}{\partial a \, \partial S} = X g'(S) f(K) < 0.$$

But the firm changes its optimal complementary investment in response to both capability and diffusion. Extending the negative interaction from current profits to the entire value function therefore requires that the endogenous investment response not reverse it. Internet Appendix Section IA.5 states sufficient conditions under which the post-deployment control problem preserves this ordering.

The model does not equate empirical average q with $\mathcal{R}$. The marginal q^K governing investment in Equation (8) is the shadow value of an additional unit of complementary capital. The empirical average-q measures also contain installed tangible and intangible capital, adjustment costs, growth options, and any scarcity value. Peters–Taylor total q improves the measured capital perimeter but does not directly decompose economic rents (Peters and Taylor, 2017; Crouzet and Eberly, 2023).

The framework yields four testable implications:

P1. Selection and timing. Greater capability lowers the deployment threshold. Its empirical counterpart is that firms with greater pre-disclosure capability should reach their first strict deployment disclosure at lower industry diffusion states.

P2. Complementary investment. Deployment is followed by investment in complementary intangible capital when marginal q^K exceeds one. Because the firm incurs investment and adjustment costs during installation, operating returns need not improve immediately.

P3. Valuation compression. Conditional on deployment, valuation should be more favorable at low diffusion. Under the maintained decreasing-differences condition, diffusion should also reduce the incremental valuation associated with pre-existing capability.

P4. Defensive deployment. Peer diffusion can increase the likelihood that firms remaining behind deploy, even as it reduces the valuation advantage associated with deployment. The first-disclosure hazard evidence provides a reduced-form counterpart of this prediction.

The model's stopping time represents the firm's unobserved internal deployment decision, whereas the data identify the first strict public disclosure providing evidence of deployment. The empirical event may therefore lag internal deployment. I interpret the diffusion gradient, the capability-by-diffusion interaction, the complementary-investment response, and the first-disclosure hazard as reduced-form counterparts of the model's comparative statics rather than as estimates of its structural parameters.

The realized operating-return analysis asks whether the valuation response is followed by observable operating gains within the available horizon. The balance-sheet results describe how firms adjust after disclosure, but the model does not assign those patterns a structural financing interpretation.

4 Empirical Design and Valuation Outcomes

The empirical analysis proceeds in four steps. First, I map each firm's first strict deployment disclosure to the relevant fiscal period and define the valuation outcomes and comparison firms. Second, I estimate the average post-disclosure valuation path using interaction-weighted event studies. Third, I examine how that path varies with the predetermined industry diffusion state using both a first-mover–follower split and a continuous-gradient specification. Finally, within first movers, I examine whether valuation is lower as industry peers subsequently catch up. Throughout, treatment is absorbing: once a firm makes a strict deployment disclosure, it remains in the treated state in subsequent periods.

4.1 Treatment timing and comparison firms

Let i index firms, t fiscal years, and $g(i)$ the firm's GICS industry. Let G_i denote the treatment-cohort year constructed from firm i's first strict deployment disclosure, with

$G_i = \infty$ for never-disclosing firms. For eventual disclosers, define event time $K_{it} = t - G_i$ and the absorbing treatment indicator

$$D_{it} = \mathbf{1}\{G_i < \infty\}\mathbf{1}\{t \geq G_i\}.$$

Unless otherwise noted, standard errors are clustered by firm. Estimator-specific joint tests, bootstrap procedures, and simultaneous confidence bands are described with the corresponding specifications.

Aligning disclosure to fiscal periods. A disclosure dated in calendar time may fall within a fiscal period, so I map each exact disclosure date to Compustat fiscal-period boundaries. For balance-sheet outcomes measured at the fiscal year-end, a firm–year is treated when its fiscal year-end falls on or after the disclosure, defining the cohort year G_i^{stock}. For flow outcomes accumulated over the fiscal year, the cohort year G_i^{flow} is the first fiscal year lying entirely after the disclosure. The fiscal period containing the disclosure is therefore excluded from the treated state for flow outcomes because it combines pre- and post-disclosure activity. Appendix A.1 provides the exact timing definitions. Unless noted otherwise, G_i refers to G_i^{stock} for stock outcomes and G_i^{flow} for flow outcomes.

Never-adopter firms. I define a never-adopter firm as one for which I never observe a strict AI deployment disclosure in any transcript through 2024, conditional on the firm having at least one transcript during the sample period. This is an observational definition: never-adopter firms have no observed strict disclosure, but are not firms known never to have deployed AI internally. The baseline retains firm–years without a transcript and classifies them as untreated until a strict disclosure is observed. Because such periods provide no opportunity to observe a disclosure, I assess sensitivity to this treatment of the untreated risk set using the current-period-observable and right-censored samples defined in Section 2.4.

4.2 Valuation outcomes

The valuation analyses use three related ratios that differ in their capital perimeter. Let MVE_{it} denote market equity, $\mathrm{DEBT}_{it} = \mathrm{DLTT}_{it} + \mathrm{DLC}_{it}$, K_{it}^{int} internally generated knowledge and organization capital, and K_{it}^{tot} total purchased and internally generated intangible capital.

Raw book-assets q. The headline dynamic outcome is

$$q_{it}^{\mathrm{book}} = \frac{\mathrm{AT}_{it} + \mathrm{MVE}_{it} - \mathrm{CEQ}_{it} - \mathrm{TXDB}_{it}}{\mathrm{AT}_{it}}, \tag{19}$$

with missing TXDB set to zero. This measure retains the broadest coverage and serves as the headline outcome for the average valuation path. The two capitalized ratios below assess whether the results depend on the measured capital perimeter.

Denominator-only intangible-adjusted q. To ask whether omitted internally generated capital mechanically drives the raw ratio, I use

$$q_{it}^{\mathrm{book,int}} = \frac{\mathrm{AT}_{it} + \mathrm{MVE}_{it} - \mathrm{CEQ}_{it} - \mathrm{TXDB}_{it}}{\mathrm{AT}_{it} + K_{it}^{\mathrm{int}}}. \tag{20}$$

The market-value numerator is unchanged while the estimated replacement value of internally generated knowledge and organization capital is added to the denominator. The measure is designed to test a specific denominator-mismeasurement explanation; it is not a direct estimator of rents.

Peters–Taylor total q. The broader operating-capital perimeter is

$$q_{it}^{\mathrm{PT}} = \frac{\mathrm{MVE}_{it} + \mathrm{DEBT}_{it} - \mathrm{ACT}_{it}}{\mathrm{PPEGT}_{it} + K_{it}^{\mathrm{tot}}}. \tag{21}$$

It scales operating value by gross physical plus total intangible replacement capital and is used as an accounting-perimeter robustness and in the capability analysis, rather than as the sole headline dynamic outcome. The reconstructed series correlates 0.976 with the

official WRDS `totalq` series on the overlapping sample.

All three ratios are winsorized within fiscal year at the first and ninety-ninth percentiles. Internet Appendix Section IA.2.2 provides the capitalization procedure, depreciation parameters, sample screens, and reconciliation details.

4.3 Average post-disclosure dynamics

I estimate dynamic changes around the first strict deployment disclosure under staggered timing using the interaction-weighted event-study estimator of Sun and Abraham (2021). Treatment is based on the first transcript with $AID = 1$ and is mapped to the relevant stock- or flow-aligned cohort defined in Section 4.1. The estimator permits treatment effects to vary across disclosure cohorts and event time and avoids the negative-weight comparisons that can arise in conventional two-way fixed-effects event studies.

Setup and specification. Let G_i denote firm i's treatment-cohort year, with $G_i = \infty$ for never-disclosing firms, and let $K_{it} = t - G_i$ denote event time for eventual disclosers. Let $\mathcal{J}_g$ denote the set of event-time indicators included for disclosure cohort g. I estimate

$$Y_{it} = \alpha_i + \lambda_t + \sum_{g} \sum_{k \in \mathcal{J}_g \setminus \{-1\}} \theta_{k,g} \, \mathbf{1}\{G_i = g\} \mathbf{1}\{K_{it} = k\} + \varepsilon_{it}, \tag{22}$$

where α_i and λ_t are firm and calendar-year fixed effects and $k = -1$ is the omitted reference period. Not-yet-disclosing and never-disclosing firms observed in the same calendar years provide the untreated comparisons.

Identification requires that, absent the disclosure-related transition, treated and comparison firms would have followed parallel outcome trends. I assess whether differential outcome movements are visible before disclosure using the lead coefficients and joint tests described below.

Fixed effects and controls. The baseline event studies include firm and calendar-year fixed effects and no time-varying controls. Excluding time-varying controls avoids conditioning on variables that may themselves respond to deployment and preserves the

coefficients as total reduced-form changes relative to the omitted event year $k = -1$.

Interaction-weighted aggregation. For a reported event time k, let $\mathcal{G}_k$ denote the supported disclosure cohorts contributing to that horizon. Following Sun and Abraham (2021), I aggregate the corresponding cohort-specific estimates as

$$\beta_k^{\mathrm{IW}} = \sum_{g \in \mathcal{G}_k} \omega_{k,g} \theta_{k,g}, \qquad \omega_{k,g} = \frac{N_{k,g}}{\sum_{g' \in \mathcal{G}_k} N_{k,g'}}, \tag{23}$$

where $N_{k,g}$ is the number of treated observations contributed by cohort g at event time k. The figures report $\widehat{\beta}_k^{\mathrm{IW}}$, 95% confidence intervals, and the corresponding treated-firm count $N(k)$.

Windows, support, and pre-trends. For the first-mover and follower paths, I define the endpoint-binned event time

$$\widetilde{K}_{it} = \max\{-5, \min(K_{it}, 5)\}$$

and use

$$\mathcal{K} = \{-5, -4, -3, -2, 0, 1, 2, 3, 4, 5\},$$

with $k = -1$ omitted. Thus, the left endpoint pools observations with $K_{it} \leq -5$, while the right endpoint pools observations with $K_{it} \geq +5$. For these group-specific specifications, $\widetilde{K}_{it}$ replaces K_{it} in Equation (22), and both paths use the same displayed window.

I report a group-specific event-time coefficient only when it is supported by at least 30 treated firms and at least three disclosure cohorts. The figures report $N(k)$ alongside each path so that the treated support at every horizon is visible.

For the principal group-specific paths, I report three pre-trend diagnostics using the firm-clustered covariance matrix. The primary local Wald test jointly tests the near-event leads $k = -4, -3, -2$. A supplementary all-lead test additionally includes the pooled $k \leq -5$ endpoint, which represents more distant pre-disclosure history and may have a different cohort composition. Finally, a one-degree-of-freedom linear-drift test examines

whether the pre-disclosure coefficients display a systematic linear trend.

4.4 Diffusion-dependent valuation

I measure diffusion heterogeneity in two complementary ways. The first-mover–follower split translates the predetermined diffusion state into an economically transparent group comparison. The continuous specification then uses the full distribution of the same state and tests whether the post-disclosure valuation path varies with diffusion without relying on the quartile cutoff.

4.4.1 Predetermined diffusion and the first-mover split

The first heterogeneity design divides strict disclosers according to the industry diffusion state prevailing before their first strict disclosure. Let $g(i)$ denote firm i's six-digit GICS industry and let G_i denote its first strict-disclosure cohort year. I define firm i's predetermined diffusion state as

$$S_i \equiv S^{-i}_{g(i),G_i-1}, \tag{24}$$

where $S^{-i}_{g(i),G_i-1}$ is the share of other firms in the same industry that had already disclosed strict deployment by the end of the preceding year. The leave-one-out construction excludes firm i from its own industry share, while the one-year lag prevents the firm's disclosure and other disclosures occurring in the same year from mechanically affecting its classification.

Among firms with a strict disclosure in industry $g(i)$, firm i is classified as a first mover when S_i lies at or below the 25th percentile of the industry's discloser-specific diffusion-state distribution. The remaining classified firms are followers. The baseline classification imposes no minimum-industry-size screen, and each firm's classification is held fixed in every discrete first-mover/follower analysis. The strict-disclosure cohort aligned to fiscal year-end outcomes contains 1,293 firms: 383 first movers and 910 followers.

Interpreting the diffusion state. The state S_i measures how common strict AI deployment disclosure had become among the firm's industry peers before its own disclosure.

When S_i is low, deployment is relatively uncommon and may provide greater differentiation. When S_i is high, deployment is closer to an industry-parity input and may provide less incremental valuation advantage.

Group-specific event-time paths. I first estimate separate interaction-weighted Sun–Abraham event studies using Equation (22). The first specification includes first movers and never-disclosing firms, and the second includes followers and never-disclosing firms. Both use the common event window $k \in [-5, +5]$ and the same never-disclosing comparison pool. These specifications describe how the valuation paths evolve separately for the two groups.

4.4.2 Direct first-mover–follower contrasts

Separately estimated confidence intervals do not provide a direct test of the difference between the two groups. I therefore estimate the first-mover-minus-follower contrast using the untreated-outcome imputation estimator of Borusyak et al. (2024). I first estimate

$$Y_{it} = \alpha_i + \lambda_t + u_{it} \qquad \text{using observations for which } D_{it} = 0, \tag{25}$$

where the untreated estimation sample contains pre-disclosure observations for eventual disclosers and all available observations for never-disclosing firms. For each treated firm–year, the imputed untreated outcome and estimated treatment effect are

$$\widehat{Y}^0_{it} = \widehat{\alpha}_i + \widehat{\lambda}_t, \qquad \widehat{\tau}_{it} = Y_{it} - \widehat{Y}^0_{it}. \tag{26}$$

For a post-disclosure window W, let

$$\mathcal{T}_{iW} = \{t : K_{it} \in W,\ D_{it} = 1,\ Y_{it} \text{ is observed}\}$$

denote the treated observations available for firm i in that window. For every firm with

$|\mathcal{T}_{iW}| \geq 1$, I first calculate the firm's average imputed treatment effect:

$$\widehat{\tau}_{iW} = \frac{1}{|\mathcal{T}_{iW}|} \sum_{t \in \mathcal{T}_{iW}} \widehat{\tau}_{it}. \tag{27}$$

This is the baseline any-observation support rule: a firm contributes when its outcome is observed in at least one year of the relevant window. Let $\mathcal{I}_{hW}$ denote the firms in group $h \in \{\mathrm{FM}, \mathrm{F}\}$ that contribute to window W, where FM denotes first movers and F denotes followers. The group-specific average and the direct contrast are

$$\widehat{\tau}_{hW} = \frac{1}{|\mathcal{I}_{hW}|} \sum_{i \in \mathcal{I}_{hW}} \widehat{\tau}_{iW}, \qquad \widehat{\Delta}_W = \widehat{\tau}_{\mathrm{FM},W} - \widehat{\tau}_{\mathrm{F},W}. \tag{28}$$

The pre-specified windows are $W = [+1, +2]$ and $W = [+4, +5]$. Inference for $\widehat{\Delta}_W$ uses a 999-replication firm-cluster bootstrap that resamples firms and re-estimates the untreated firm and calendar-year effects in every replication.

Figure 3 reports the group-specific Sun–Abraham paths. Panel B of Figure 2 and Table 2 report the corresponding direct window contrasts from Equation (28). Because the event-time paths and the window contrasts use different estimators, the averages of the plotted coefficients need not equal the imputed window estimates exactly.

Support robustness. Internet Appendix Section IA.6.1 and Table IA.9 report two checks on the first-mover–follower comparison. First, I repeat the classification after restricting the sample to industries with at least eight classified disclosers, so that the within-industry diffusion quartiles are not formed from very small samples. Second, I replace the baseline rule that requires an outcome in at least one year of a two-year window with a complete-window rule requiring the outcome in both years. Under both checks, the first-mover-minus-follower difference in years $[+1, +2]$ remains positive and statistically significant across all three valuation measures, while the differences in years $[+4, +5]$ remain statistically indistinguishable from zero.

Observation and disclosure-timing robustness. Among eventual disclosers that are calendar-eligible to contribute at a given horizon, sample retention through event time $k = +7$ exceeds 95% for both first movers and followers, with little differential retention between the groups (Table A.3). The medium-run comparison is therefore unlikely to be generated by differential survivorship.

I also re-estimate the first-mover-minus-follower contrast using the current-period-observable and right-censored samples defined in Section 2.4. A separate observed-disclosure-gap specification retains all never-disclosing firms but restricts eventual disclosers to firms whose interval between the last observed non-strict transcript and the first strict disclosure is no more than 365 days. This restriction sharpens the dating of the transition in public evidence. Internet Appendix Section IA.6 reports these observation and disclosure-timing robustness results.

4.4.3 Continuous diffusion gradient

The first-mover split summarizes the diffusion environment using a single cutoff. A complementary specification uses the full predetermined diffusion state and allows the event-time path to vary continuously with it. The predetermined state avoids mechanical same-year feedback but may still coincide with industry-specific technology or demand shocks. The continuous specification therefore includes industry-by-year fixed effects, which absorb shocks common to firms in the same industry and calendar year.

Design. Let

$$S_i = S^{-i}_{g(i),G_i-1}$$

be the frozen, firm-level diffusion state defined in Section 4.4.1, and let $\bar{S}$ be its mean among eventual disclosers. For eventual disclosers, define the endpoint-binned event time

$$\widetilde{K}_{it} = \max\{-5, \min(K_{it}, 5)\},$$

and let

$$\mathcal{K} = \{-5, -4, -3, -2, 0, 1, 2, 3, 4, 5\},$$

with $k = -1$ omitted. I pool eventual and never-disclosing firms and estimate

$$Y_{it} = \alpha_i + \lambda_{g(i),t} + \sum_{k \in \mathcal{K}} \beta_k \, \mathbf{1}\{\widetilde{K}_{it} = k\} + \sum_{k \in \mathcal{K}} \delta_k \, (S_i - \bar{S}) \mathbf{1}\{\widetilde{K}_{it} = k\} + \varepsilon_{it}. \tag{29}$$

For never-disclosing firms, all event-time indicators and their interactions are set to zero, so these firms remain in the regression as untreated comparison observations. The diffusion state enters only through the event-time interactions for eventual disclosers; a separate level term for S_i would be absorbed by the firm fixed effects.

The specification includes firm fixed effects α_i and industry-by-year fixed effects $\lambda_{g(i),t}$. The latter absorb technology, demand, and other shocks common to firms in the same industry and calendar year. Unlike the cohort-specific Sun–Abraham paths, this is a pooled event-time interaction specification and does not apply the Sun–Abraham treated-firm or cohort-support screens. It uses all observations with a nonmissing outcome and reports the full endpoint-binned path.

The coefficients $\{\beta_k\}$ describe the fitted event-time path at the eventual-discloser mean diffusion state $\bar{S}$. The coefficients $\{\delta_k\}$ measure how that path changes with the predetermined diffusion state. A negative post-disclosure δ_k means that the valuation path is lower among firms disclosing when deployment is already more widespread in their industry.

Inference on the diffusion gradient. I first use the firm-clustered covariance matrix to test

$$H_0 : \delta_0 = \delta_1 = \cdots = \delta_5 = 0.$$

This joint Wald test asks whether the post-disclosure event-time path varies with the predetermined diffusion state.

I then illustrate the magnitude and timing of this heterogeneity by evaluating the fitted path at the 25th percentile of diffusion among eventual disclosers, S_{low}, and at the

75th percentile, S_{high}. At event time k, their low-minus-high difference is

$$\widehat{\delta}_k \left(S_{\text{low}} - S_{\text{high}} \right).$$

I plot these contrasts with simulation-based simultaneous 95% confidence bands derived from the joint covariance matrix of the interaction coefficients. The bands account for examining the entire event-time path rather than treating each horizon as an independent pointwise comparison.

4.5 Peer catch-up within first movers

The preceding designs compare firms disclosing at different predetermined diffusion states. As a complementary within-first-mover test, I next examine whether valuation is lower when peer diffusion subsequently rises. For first mover i, define

$$\text{catchup}_{it} = S^{-i}_{g(i),t} - S_i,$$

the current leave-one-out industry diffusion share minus the firm's predetermined diffusion state at disclosure. I estimate, over event years $k \in [1, 5]$,

$$Y_{it} = \alpha_i + \lambda_t + \phi_{K_{it}} + \psi \ \text{catchup}_{it} + \varepsilon_{it},$$

using only first movers, with firm, calendar-year, and event-time fixed effects and firm-clustered standard errors. The event-time fixed effects absorb the average post-disclosure path, so ψ compares first movers observed at the same event-time horizon but exposed to different amounts of subsequent peer diffusion. A negative ψ indicates that greater peer catch-up is associated with lower valuation conditional on these fixed effects.

First-mover status remains fixed at the predetermined classification and is never recomputed using contemporaneous diffusion. Panel D of Figure 2 reports the pooled estimate of ψ over event years +1 through +5 for raw book-assets q, intangible-adjusted q, and Peters–Taylor total q, scaled to a ten-percentage-point increase in peer diffusion.

Because peer catch-up may coincide with other industry developments, I interpret ψ as complementary within-first-mover evidence.

5 Diffusion and the Value of AI Deployment

This section establishes the paper's central empirical result: the valuation response to disclosed AI deployment is larger and arrives earlier when deployment remains uncommon among industry peers. I first document the average post-disclosure valuation path. The first-mover–follower comparison then provides the headline economic magnitudes, the continuous gradient tests whether the result depends on the quartile cutoff, and the peer-catch-up specification examines whether the same ordering appears as diffusion subsequently increases. Section 6 uses capitalization, capability, and realized operating outcomes to interpret this diffusion-dependent valuation pattern.

5.1 Average post-disclosure valuation

The primary outcome is raw book-assets q, defined in Section 4.2. Panel A of Figure 2 plots the interaction-weighted Sun–Abraham path for the average adopter with firm and industry-by-year fixed effects. Valuation is flat before disclosure and rises over the medium run, reaching approximately +0.30 at $k = 4$ and +0.38 at $k = 5$, relative to $k = -1$. In the baseline specification with firm and calendar-year fixed effects, the corresponding estimates are +0.24 and +0.33. The similar medium-run path across the two fixed-effect structures shows that the average revaluation is not absorbed by shocks common to firms in an industry-year. This is the average response across adopters; the paper's distinctive evidence concerns how that response varies with the predetermined diffusion state.

5.2 First movers versus followers

The comparison uses the discrete split defined in Section 4.4.1: within each industry, a first mover is an adopter whose predetermined diffusion state $S_i = S^{-i}_{g(i),G_i-1}$ lies at or below the industry's 25th percentile, so first movers deploy when fewer than about fifteen

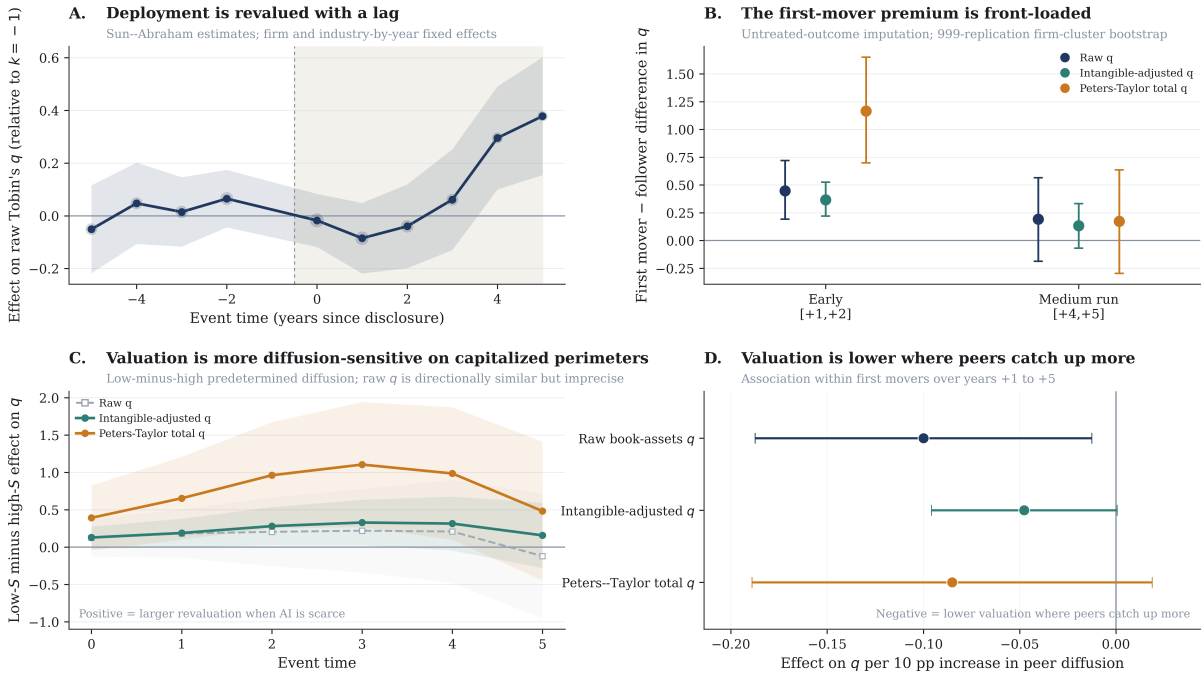


Figure 2: **Diffusion and the early-mover valuation premium.** Panel A reports the average-adopter interaction-weighted Sun–Abraham path for raw book-assets q, with firm and industry-by-year fixed effects. Panel B reports first-mover-minus-follower differences over the pre-specified $[+1, +2]$ and $[+4, +5]$ windows using the untreated-outcome imputation estimator. Panel C reports fitted low-minus-high predetermined-diffusion contrasts from the continuous event-time specification; bands are simultaneous 95% sup-t intervals. Panel D reports the pooled association between subsequent peer catch-up and each valuation measure within first movers over event years +1 through +5, expressed per ten-percentage-point increase in peer diffusion. Event time is aligned to the first fiscal year-end on or after disclosure, and $k = -1$ is omitted. Panel B inference uses a 999-replication firm-cluster bootstrap; the remaining standard errors are clustered by firm.

percent of their peers have, compared with about thirty percent for the typical adopter. The discrete estimation sample contains 383 first movers and 910 followers, and both groups span the sample period because the split is based on diffusion at disclosure rather than calendar time.

Figure 3 reports separate interaction-weighted Sun–Abraham paths for first movers and followers against the common never-disclosing comparison pool for raw book-assets q, intangible-adjusted q, and Peters–Taylor total q. Panel B of Figure 2 and Table 2 report the direct first-mover-minus-follower contrasts over the pre-specified $[+1, +2]$ and $[+4, +5]$ windows.

The raw-q components clarify the source of the early differential. Over $[+1, +2]$, the imputed effect is $+0.177$ for first movers and -0.270 for followers, producing a first-mover-minus-follower difference of $+0.447$ ($t = 3.39$, bootstrap $p = 0.001$). Thus the early contrast reflects both an earlier positive revaluation among first movers and an initial relative decline among followers. By $[+4, +5]$, both groups have positive estimated effects—$+0.434$ for first movers and $+0.242$ for followers—and the difference has narrowed to $+0.192$ ($t = 1.00$, $p = 0.321$). The result is therefore a front-loaded difference in the timing of revaluation, rather than evidence that followers never revalue.

Figure 3 shows an analogous qualitative pattern on the two capitalized perimeters: first movers revalue earlier, while followers initially decline relative to the never-discloser counterfactual and subsequently recover. The formal early-window first-mover-minus-follower difference is $+0.367$ on intangible-adjusted q ($t = 4.73$, $p = 0.001$) and $+1.165$ on Peters–Taylor total q ($t = 4.80$, $p = 0.001$). By $[+4, +5]$, the corresponding differences have attenuated to $+0.133$ ($p = 0.220$) and $+0.173$ ($p = 0.487$). Section 6.1 discusses what this stability across accounting perimeters does and does not imply.

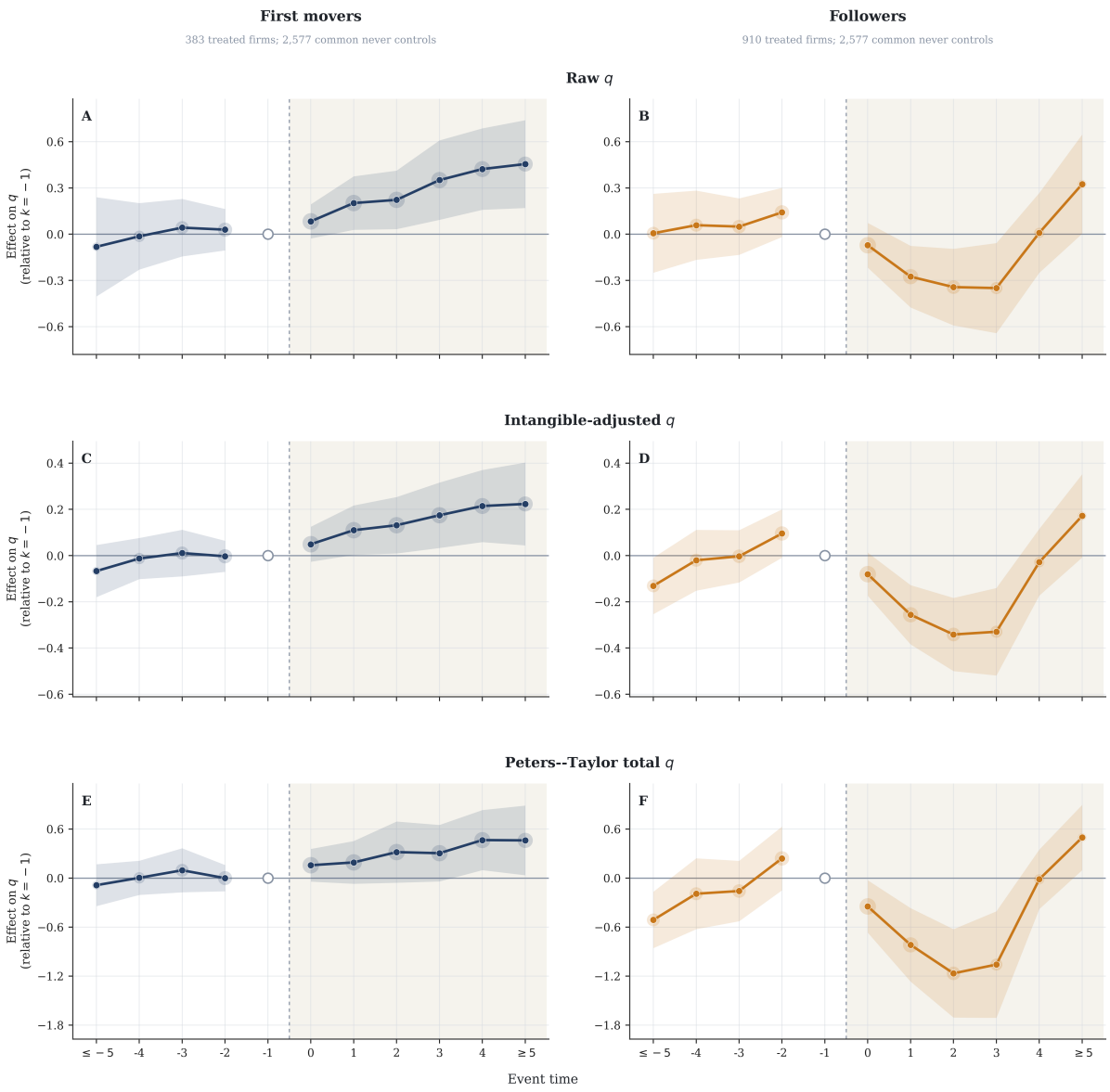


Figure 3: **First-mover and follower valuation paths across accounting perimeters.** The figure reports interaction-weighted Sun–Abraham event-study estimates separately for first movers and followers against the common never-discloser pool. The rows report raw book-assets q, denominator-adjusted q, and Peters–Taylor total q. First-mover status is fixed using the predetermined leave-one-out diffusion state at the firm's first strict deployment disclosure. Event time is aligned to the first fiscal year-end on or after disclosure; $k = -1$ is omitted, and endpoints pool $k \leq -5$ and $k \geq 5$. Shaded regions are 95% confidence intervals based on firm-clustered standard errors. The vertical scales differ across rows because the three valuation ratios use different capital and firm-value perimeters. The early differential reflects both an earlier positive response among first movers and an initial relative decline among followers; followers subsequently recover at later supported horizons. Formal inference on the first-mover-minus-follower difference is reported in Table 2.

Table 2: First-mover minus follower difference in the q revaluation

Window	Raw q	Intangible-adjusted q	Peters–Taylor total q
$[+1, +2]$	$+0.447^{***}$	$+0.367^{***}$	$+1.165^{***}$
	$(t = 3.39)$	$(t = 4.73)$	$(t = 4.80)$
$[+4, +5]$	$+0.192$	$+0.133$	$+0.173$
	$(t = 1.00)$	$(t = 1.25)$	$(t = 0.72)$

Notes: First-mover-minus-follower difference in the imputed treatment effect, averaged over each window, from the untreated-outcome imputation estimator. Inference uses a 999-replication firm-cluster bootstrap that re-estimates the untreated firm and year effects in every draw; all 999 draws were successful. Bootstrap p-values are 0.001, 0.001, and 0.001 in the early window and 0.321, 0.220, and 0.487 in the later window. First movers and followers are the predetermined diffusion split of Section 4.4.1. The discrete estimation sample contains 383 first movers and 910 followers. Outcome-specific support is 382/761 first movers/followers for raw and intangible-adjusted q at $[+1, +2]$, and 342/687 for Peters–Taylor total q; support falls at later horizons. The $[+6, +8]$ window is omitted because only 84 followers contribute to raw and adjusted q and 75 to total q; all three estimates are insignificant in the Internet Appendix. Raw q is the headline; the adjusted columns are the mismeasurement robustness of Section 6.1. $^{***}p < 0.01$.

5.3 Continuous diffusion gradient

The continuous specification tests whether the same diffusion ordering appears over the full distribution of S_i, rather than depending on the 25th-percentile cutoff used to define first movers. Panel C of Figure 2 reports the specification defined in Section 4.4.3. I evaluate the fitted path at the 25th and 75th percentiles of the predetermined diffusion state and report the low-diffusion path minus the high-diffusion path. A positive contrast therefore means that the fitted post-disclosure valuation response is larger for firms disclosing at low predetermined diffusion. This is the cutoff-free counterpart of the first-mover–follower comparison.

The low-minus-high-diffusion contrast is positive over the post-disclosure horizon on the capitalized measures, reaching roughly one point of Peters–Taylor q by the third year. The joint null that the post-disclosure path does not vary with diffusion is rejected for intangible-adjusted q ($p = 0.037$) and Peters–Taylor total q ($p = 0.006$). On raw q, the point estimates have the same sign but are imprecise and the joint test does not reject ($p = 0.40$). The continuous-gradient evidence is therefore carried primarily by the capitalized measures, while raw q provides directional corroboration. Panel C uses

simultaneous sup-t bands so that the horizons are interpreted jointly. The continuous design and the discrete first-mover contrast are different summaries of the same predetermined diffusion state and point in the same economic direction.

Internet Appendix Figure IA.14 decomposes these contrasts into the underlying fitted paths. On the two intangible-capitalized perimeters, the low-diffusion path remains close to zero over the first several post-disclosure years, whereas the high-diffusion path becomes substantially negative. The significant continuous gradient therefore reflects primarily an unfavorable valuation path among firms disclosing once deployment is already common, rather than a uniformly large positive response among low-diffusion firms. Raw book-assets q displays the same ordering through most of the post-disclosure horizon, but with wider intervals and no rejection of the joint interaction test.

The capitalized diffusion gradient is not driven by the inclusion of financial firms or utilities. On the nonfinancial, nonutility sample, the joint post-disclosure null remains unrejected for raw book-assets q ($p = 0.272$), but is rejected for intangible-adjusted q ($p = 0.024$). On the operating-firm Peters–Taylor sample, which also requires a meaningful physical-plus-intangible replacement-capital denominator, the joint null is rejected at $p = 0.003$. The post-disclosure interaction coefficients on both capitalized measures are negative from event year zero through year four. Full sample definitions and estimates are reported in Internet Appendix Section IA.9.2.

5.4 Peer catch-up within first movers

Panel D of Figure 2 reports the within-first-mover specification defined in Section 4.5. The coefficient measures the association between subsequent peer diffusion and valuation over event years +1 through +5, conditional on firm, calendar-year, and event-time fixed effects.

A ten-percentage-point increase in peer diffusion is associated with approximately 0.10 lower raw book-assets q ($p = 0.025$) and 0.05 lower intangible-adjusted q ($p = 0.050$). The corresponding Peters–Taylor total-q estimate is also negative, at approximately -0.09, but is imprecisely estimated. Conditional on firm, calendar-year, and event-time fixed

effects, greater peer diffusion is therefore associated with lower raw and intangible-adjusted valuation among first movers, while the total-q estimate points in the same direction.

5.5 Pre-disclosure diagnostics and event validation

Pre-disclosure dynamics. For the group-specific event studies, the relevant diagnostics are the pre-disclosure paths of first movers and followers relative to the common never-disclosing comparison pool. On raw book-assets q, the local joint test of the near-event leads does not reject for first movers ($p = 0.92$) or followers ($p = 0.37$), and neither group displays a significant linear pre-trend. The local and linear-drift tests are also clean for both groups on the two capitalized measures. The all-lead tests reject for followers on intangible-adjusted and Peters–Taylor total q, but those rejections are driven by the pooled distant endpoint rather than by the near-event leads (Table C.1, Panel A). The near-event diagnostics therefore do not indicate a differential trend immediately preceding disclosure.

Deployment-stage validation. The valuation pattern is concentrated in disclosures indicating deployment rather than exploration or generic AI discussion. When firms are sorted by the intensity of their AI activity, the medium-run revaluation is largest for confirmed deployment, weaker for pilot-stage exploration, and approximately flat for firms that only mention AI, all relative to the same no-AI-mention comparison pool. Internet Appendix Section IA.7 reports these estimates.

Short-window market response. First strict deployment disclosures are also economically meaningful dated market events. Mean Fama–French five-factor-adjusted CAR$[0, +1]$ is 0.56% for first movers and 0.44% for followers. The regression-adjusted difference between the two groups is positive but imprecisely estimated ($p = 0.24$). I use these estimates to establish the market relevance of the disclosure date, rather than as separate identification of the medium-run diffusion effect. Internet Appendix Section IA.8 reports the complete announcement-return results and power calculations.

6 What Is Being Priced? Capitalization, Capability, and Realized Returns

Section 5 establishes that the valuation response to disclosed deployment declines with the predetermined diffusion state. This section evaluates three interpretations of that pattern. First, I ask whether omitted internally generated capital mechanically produces the early valuation differential. Second, I test whether the valuation slope on pre-disclosure capability declines as diffusion increases. Third, I examine whether the early valuation differential is followed by observable operating outperformance.

6.1 Capitalization and denominator mismeasurement

The most serious mechanical alternative is that book assets omit the software, data, and organizational know-how accumulated around deployment. If the market capitalizes those assets while accounting continues to expense them, raw book-assets q can rise even when the firm earns a competitive return on the installed capital. The relevant question is therefore not whether adjusted average q remains above one, but whether the early first-mover-minus-follower differential is absorbed when the preferred omitted capital stock enters the measured asset base.

The denominator-only adjustment and Peters–Taylor total q are defined in Section 4.2, with full construction details in Internet Appendix Section IA.2.2. As Table 2 shows, the early first-mover-minus-follower difference remains positive and statistically significant on both capitalized valuation measures, while the later-window differences are substantially smaller and statistically indistinguishable from zero. Capitalizing internally generated intangibles therefore changes the scale of the estimates but does not absorb their early concentration.

The survival of the differential after capitalization indicates that omitted intangible capital does not provide a complete mechanical explanation. Because average q also contains growth options, adjustment costs, and any scarcity value, the result is interpreted jointly with the diffusion and capability evidence rather than as a stand-alone

decomposition of rents (Crouzet and Eberly, 2023; Peters and Taylor, 2017).

6.1.1 Parameter sensitivity: the intangible-capital grid

The intangible adjustment depends on three parameters: the fraction of SG&A treated as organizational investment (λ), and the depreciation rates of organizational capital (δ_O) and knowledge capital (δ_R). Because the whole point of the adjustment is to net out intangibles, the result should not hinge on one parameterization. Table 3 re-estimates the early first-mover differential on Peters–Taylor total q under the published baseline and six perturbations. The estimate ranges from +1.06 to +1.29 and is significant at the one-percent level in every case. The survival of the early differential is therefore not specific to one intangible-capital calibration.

Table 3: First-mover differential across the intangible-capital parameter grid (Peters–Taylor total q, $[+1, +2]$)

Intangible-capital parameterization	First mover − follower	t
Baseline ($\lambda = 0.30,\ \delta_R = 0.15,\ \delta_O = 0.20$)	$+1.165^{***}$	4.80
$\lambda = 0.20$	$+1.286^{***}$	4.70
$\lambda = 0.40$	$+1.062^{***}$	4.90
$\delta_O = 0.15$	$+1.144^{***}$	4.86
$\delta_O = 0.33$	$+1.201^{***}$	4.69
$\delta_R = 0.20$	$+1.210^{***}$	4.86
$\delta_R = 0.30$	$+1.287^{***}$	4.91

Notes: Each row rebuilds the internal intangible-capital stock under the indicated parameters, recomputes Peters–Taylor total q, and re-estimates the first-mover-minus-follower difference over $[+1, +2]$ with the imputation estimator. Inference uses 999 firm-cluster bootstrap replications with the untreated fixed effects re-estimated in every draw; all draws were successful and every centered-bootstrap p-value equals 0.001. $^{***}p < 0.01$.

The same robustness holds for the whole first-mover life cycle, not just the early-window premium: Internet Appendix Section IA.9 re-estimates the first-mover-minus-follower difference at both windows across the raw, intangible-adjusted, and Peters–Taylor perimeters and all six calibrations, and capitalizing intangibles changes the scale of the premium but not its early concentration or its subsequent attenuation.

6.2 Capability and diffusion

A temporary scarcity interpretation predicts that the valuation slope on pre-disclosure capability should decline as deployment becomes common. I test this prediction in adopter-level cross-sections. For each firm, I define long-run revaluation as average q over event years $k \in \{3, 4, 5\}$ minus average q over $k \in \{-2, -1\}$ and regress it on the predetermined diffusion state, a standardized pre-disclosure capability proxy, and their interaction. All specifications include adoption-cohort fixed effects and pre-disclosure log assets.

Table 4 reports the Peters–Taylor total-q results under both sector and fine-industry fixed effects. With sector fixed effects, the capability-by-diffusion interaction is -0.386 ($t = -3.20$) for constructed intangible capital and -0.583 ($t = -4.96$) for R&D intensity. Thus, the valuation slope on capability is materially smaller when firms disclose deployment at a higher predetermined diffusion state.

The interaction remains negative but attenuates under fine-industry fixed effects, falling to -0.159 ($t = -1.19$) for constructed intangible capital and -0.200 ($t = -1.37$) for R&D intensity. The comparison shows that the capability-compression result is strongest in the variation across fine industries within broad sectors and becomes less precisely estimated when fine-industry effects absorb that variation. Internet Appendix Section IA.11 reports the complete set of valuation perimeters, capability proxies, and capitalization calibrations.

Table 4: Capability and predetermined diffusion

	Sector fixed effects			Fine-industry fixed effects		
Capability proxy × diffusion	Coef.	t	N	Coef.	t	N
Intangible capital / assets	-0.386^{***}	-3.20	634	-0.159	-1.19	634
R&D / assets	-0.583^{***}	-4.96	458	-0.200	-1.37	458

Notes: The table reports the coefficient on $S^{-i}_{g(i),G_i-1} \times A_i$ from adopter-level regressions of long-run Peters–Taylor total-q revaluation on the predetermined diffusion state, a standardized pre-disclosure capability proxy, and their interaction. Long-run revaluation is average q over $k \in \{3, 4, 5\}$ minus average q over $k \in \{-2, -1\}$. All specifications include adoption-cohort fixed effects and pre-disclosure log assets. Continuous regressors are standardized, and standard errors are clustered by industry. Sector fixed effects preserve variation across fine industries within sectors; fine-industry fixed effects absorb much of that variation and are reported as a more demanding sensitivity. $^{***}p < 0.01$.

6.3 Realized operating returns

If the valuation differential reflects future operating gains, those gains may eventually appear in fundamentals. Because complementary investment is expensed, accounting returns can decline during implementation even when expected firm value rises, as in the productivity J-curve of Brynjolfsson et al. (2021). I therefore estimate the first-mover-minus-follower difference for a pre-specified family of six operating outcomes using the same untreated-outcome imputation framework and a 999-replication firm-cluster bootstrap.

No outcome shows first-mover operating outperformance within the observed horizon. In the unbalanced specification, intangible-adjusted return on invested capital is 1.185 percentage points lower for first movers in the early window ($t = -1.84$, bootstrap $p = 0.065$). The estimate is statistically significant under inverse-probability-of-censoring weighting (-1.493 points, $p = 0.046$), but not in the balanced sample (-0.921 points, $p = 0.171$). Operating cash flow over assets is also lower, but imprecisely estimated. The complete six-outcome family and the balanced-support and IPCW sensitivities appear in Internet Appendix Section IA.12.

These results show that the valuation differential precedes observable first-mover operating outperformance. Its diffusion-sensitive interpretation comes from the capitalization, continuous-gradient, capability, and peer-catch-up evidence rather than from the realization estimates alone.

Table 5: Realized operating returns: first mover minus follower

Outcome (first mover − follower, percentage points)	$[+1, +2]$	$[+4, +5]$
Intangible-adjusted ROIC	-1.18	-0.76
	$(t = -1.84)$	$(t = -0.79)$
Operating cash flow / assets	-0.55	-1.39
	$(t = -0.45)$	$(t = -1.01)$

Notes: First-mover-minus-follower difference in the imputed treatment effect from the unbalanced specification. Inference uses a 999-replication firm-cluster bootstrap that re-estimates untreated firm and year effects in every draw; all 999 draws were successful for every outcome. Centered-bootstrap p-values for intangible-adjusted ROIC are 0.065 and 0.417, and those for operating cash flow/assets are 0.655 and 0.300. For the early adjusted-ROIC estimate, the balanced-sample result is -0.92 ($p = 0.171$) and the IPCW result is -1.49 ($p = 0.046$). First-mover status is the predetermined classification of Section 4.4.1, while event time uses the first fully post-disclosure fiscal period for these annual flow outcomes. The full six-outcome battery, including balanced and IPCW diagnostics, is in Section IA.12.

Labor-share adjustment. Internet Appendix Section IA.13 reports complementary labor-share evidence. The later-window labor-share decline is larger for followers in two of the three firm-level constructions, pointing to greater labor-saving adjustment among later adopters. A detailed analysis of the underlying workforce mechanism is left to the companion labor-market paper.

7 Complementary Investment and Financing

The valuation evidence is accompanied by two forms of post-disclosure adjustment: a shift toward intangible investment and a diffusion-dependent balance-sheet response. I document these in turn. Internet Appendix Section IA.14 reports the corresponding constructed-capital, equity-issuance, and secured-debt diagnostics.

7.1 Complementary intangible investment

Figure 4 reports three asset-scaled investment outcomes. The post-disclosure investment mix shifts toward intangible rather than physical capital. R&D expenditure over assets rises, capital expenditure over assets is flat to declining, and booked intangibles over assets increase at later horizons. R&D provides the cleanest dynamic path, while the booked-intangibles series displays some pre-disclosure movement. A constructed intangible-capital

stock provides a complementary stock measure and is reported in Internet Appendix Figure IA.18.

Internet Appendix Section IA.15 reports complementary cost-share and markup evidence. COGS and SG&A shares rise at several post-disclosure horizons, while measured production-based markups show no sustained increase among first movers. Together with the increase in R&D expenditure, this pattern is consistent with an expensed installation program rather than an immediate increase in measured pricing power.

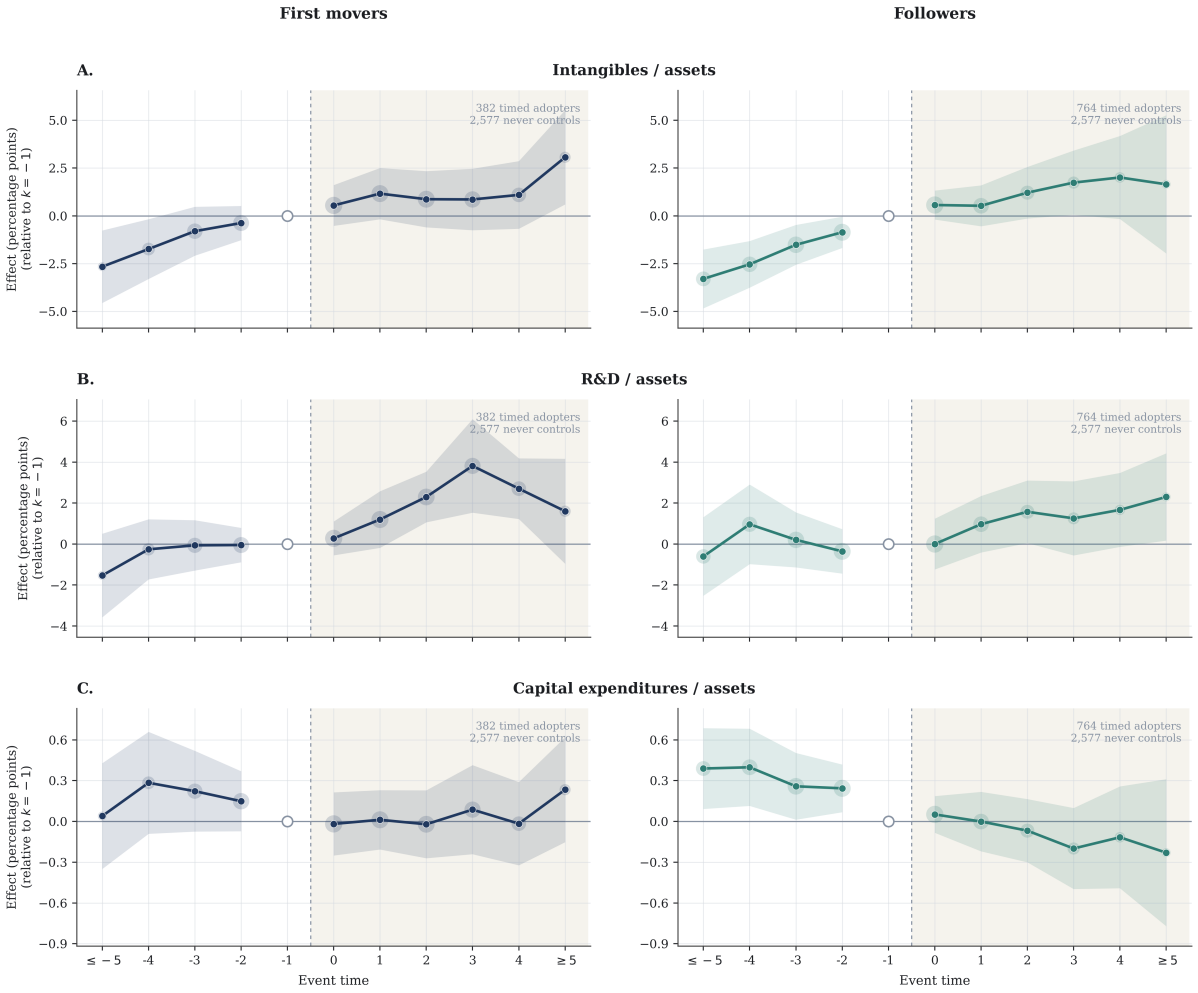


Figure 4: **Investment composition after disclosed AI deployment.** The figure reports interaction-weighted Sun–Abraham event-study estimates for booked intangibles over assets (Panel A), R&D expenditure over assets (Panel B), and capital expenditure over assets (Panel C), separately for first movers and followers against the common never-discloser pool. First-mover status uses the predetermined diffusion classification, while event time uses the first fiscal period lying entirely after the strict deployment disclosure. Event time $k = -1$ is omitted, endpoint coefficients pool $k \leq -5$ and $k \geq 5$, shaded regions are 95% confidence intervals, and marker size reflects treated-firm support. R&D is the cleanest measure of complementary intangible investment; the booked-intangibles path exhibits some pre-disclosure movement.

7.2 Debt, liquidity, and financing flows

The balance-sheet response associated with deployment differs with the diffusion position at which the firm enters. Figure 5 reports debt over assets and cash over assets. First movers draw down cash but exhibit only a modest increase in debt, whereas followers increase debt progressively after disclosure.

The direct first-mover-minus-follower contrasts confirm a medium-run divergence in debt but not in cash. Over $[+4, +5]$, the debt-over-assets response of first movers is 4.05 percentage points smaller than that of followers ($p = 0.038$). The corresponding early-window difference is -1.46 percentage points and is not statistically significant ($p = 0.236$). The cash-over-assets differences are close to zero in the early window and statistically insignificant in both windows.

Internet Appendix Table IA.18 further decomposes the debt response into secured debt over assets and the secured share of total debt. The medium-run first-mover-minus-follower difference in secured debt over assets is -2.24 percentage points ($p = 0.066$), while the secured share of total debt does not differ significantly between the two groups.

Debt and cash are balance-sheet stocks measured at the fiscal year-end. To examine financing flows, Internet Appendix Section IA.14.2 uses net external equity issuance from the cash-flow statement and aligns treatment to the first fiscal period lying entirely after disclosure. Neither group uses net external equity issuance to finance the post-disclosure adjustment. Over $[+1, +2]$, first movers repurchase 6.73 percent of beginning assets, compared with 21.05 percent among followers, producing a first-mover-minus-follower difference of $+14.32$ percentage points ($p = 0.001$). The difference disappears by $[+4, +5]$.

The equity-flow evidence shows that external equity issuance is not the post-disclosure financing margin. Followers combine higher debt with larger repurchases in the early window, indicating that the debt increase is part of a broader capital-structure reallocation.

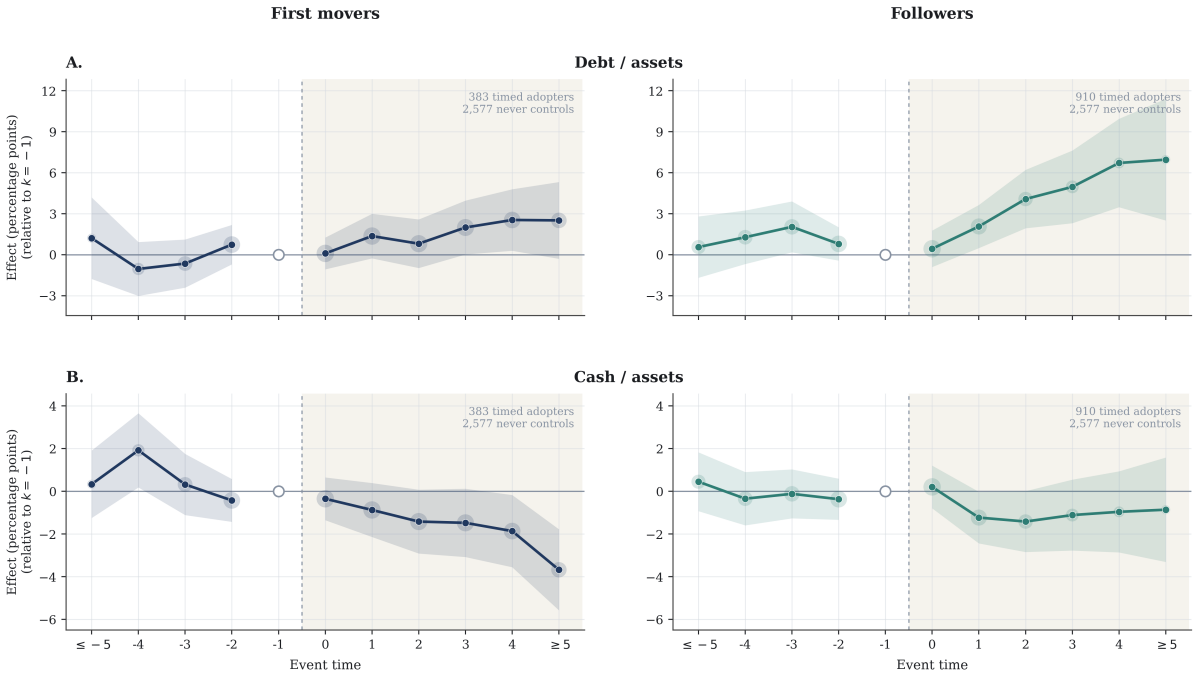


Figure 5: **Debt and liquidity after disclosed AI deployment.** The figure reports interaction-weighted Sun–Abraham event-study estimates for debt over assets (Panel A) and cash over assets (Panel B), separately for first movers and followers against the common never-discloser pool. First-mover status uses the predetermined diffusion classification. Because these are fiscal-year-end balance-sheet stocks, event time is aligned to the first fiscal year-end on or after the strict deployment disclosure. Event time $k = -1$ is omitted, endpoint coefficients pool $k \leq -5$ and $k \geq 5$, and shaded regions are 95% confidence intervals based on firm-clustered standard errors.

8 Conclusion

This paper shows that the market value associated with disclosed AI deployment depends on the industry diffusion state at which a firm first discloses deployment. I identify the first strict public disclosure of deployment in corporate transcripts, align the exact disclosure date to fiscal periods, and locate each disclosing firm on a predetermined, leave-one-out industry diffusion curve.

The central result is a difference in the timing of revaluation. The first-mover-minus-follower difference in raw book-assets q is +0.447 in the first two post-disclosure years and narrows to +0.192, statistically indistinguishable from zero, in years four and five. The early difference combines a positive response among first movers with an initial decline among followers; by the later window, both groups have positive estimated effects. Firms disclosing while deployment remains scarce therefore revalue earlier, rather than maintaining a permanently higher valuation level. The continuous diffusion gradient and the decline in first-mover valuation as peers subsequently catch up reinforce the same timing pattern.

The early differential survives capitalization of internally generated knowledge and organization capital. Pre-disclosure capability also carries a steeper valuation slope when diffusion is low in the baseline specification. After disclosure, the investment mix shifts toward R&D and booked intangibles rather than physical capital, yet no outcome in the pre-specified operating family shows first-mover outperformance within the observed horizon. Together, these findings are consistent with capability-intensive adjustment whose value is reflected in market valuation before superior operating performance becomes visible in the accounting data.

The broader implication extends beyond AI. When effective use of a general-purpose technology requires firm-specific complementary capability, diffusion changes the economic role of adoption: deployment can differentiate a firm while implementation capability remains scarce and become necessary for competitive parity once peers catch up. The real-options framework formalizes these opposing forces: diffusion raises the competitive cost of remaining behind while compressing the scarcity value of early implementation

capability. The data date public disclosure rather than latent internal deployment, and the observed valuation response does not distinguish whether the priced scarcity component is ultimately realized through operating rents or longer-duration growth opportunities. The central result is that either source of value depends systematically on where the firm enters the competitive diffusion curve. Competitive diffusion is therefore a central state variable for valuing technology adoption.

References

Babina, T., Fedyk, A., He, A. X., and Hodson, J. (2023). Firm investments in artificial intelligence technologies and changes in workforce composition. Working Paper 31325, National Bureau of Economic Research.

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

Bloom, N., Sadun, R., and Van Reenen, J. (2012). Americans do it better: Us multinationals and the productivity miracle. *American Economic Review*, 102(1):167–201.

Bonney, K., Breaux, C. L., Dinlersoz, E., Foster, L. S., Haltiwanger, J. C., and Pande, A. A. (2026). The microstructure of ai diffusion: Evidence from firms, business functions, and worker tasks. Working Paper 35141, National Bureau of Economic Research.

Borusyak, K., Jaravel, X., and Spiess, J. (2024). Revisiting event-study designs: Robust and efficient estimation. *Review of Economic Studies*, 91(6):3253–3285.

Bresnahan, T. F., Brynjolfsson, E., and Hitt, L. M. (2002). Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *Quarterly Journal of Economics*, 117(1):339–376.

Brynjolfsson, E., Li, D., and Raymond, L. R. (2025). Generative AI at work. *The Quarterly Journal of Economics*, 140(2):889–942.

Brynjolfsson, E., Rock, D., and Syverson, C. (2021). The productivity J-curve: How intangibles complement general purpose technologies. *American Economic Journal: Macroeconomics*, 13(1):333–372.

Corrado, C., Criscuolo, C., Haskel, J., Himbert, A., and Jona-Lasinio, C. (2021). New evidence on intangibles, diffusion and productivity. OECD Science, Technology and Industry Working Papers 2021/10, OECD Publishing, Paris.

Crouzet, N. and Eberly, J. C. (2023). Rents and intangible capital: A Q+ framework. *The Journal of Finance*, 78(4):1873–1916.

David, P. A. (1990). The dynamo and the computer: An historical perspective on the modern productivity paradox. *The American Economic Review*, 80(2):355–361.

Eisfeldt, A. L. and Papanikolaou, D. (2013). Organization capital and the cross-section of expected returns. *Journal of Finance*, 68(4):1365–1406.

Eisfeldt, A. L. and Papanikolaou, D. (2014). The value and ownership of intangible capital. *American Economic Review*, 104(5):189–194.

Eisfeldt, A. L., Schubert, G., and Zhang, M. B. (2023). Generative AI and firm values. Working Paper 31222, National Bureau of Economic Research.

Eloundou, T., Manning, S., Mishkin, P., and Rock, D. (2023). GPTs are GPTs: An early look at the labor market impact potential of large language models.

Fudenberg, D. and Tirole, J. (1985). Preemption and rent equalization in the adoption of new technology. *The Review of Economic Studies*, 52(3):383–401.

Grenadier, S. R. (1996). The strategic exercise of options: Development cascades and overbuilding in real estate markets. *Journal of Finance*, 51(5):1653–1679.

Lieberman, M. B. and Montgomery, D. B. (1988). First-mover advantages. *Strategic Management Journal*, 9(S1):41–58.

McElheran, K., Li, J. F., Brynjolfsson, E., Kroff, Z., Dinlersoz, E., Foster, L., and Zolas, N. (2024). AI adoption in america: Who, what, and where. *Journal of Economics & Management Strategy*, 33(2):375–415.

Peters, R. H. and Taylor, L. A. (2017). Intangible capital and the investment-q relation. *Journal of Financial Economics*, 123(2):251–272.

Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199.

Yotzov, I., Barrero, J. M., Bloom, N., Bunn, P., Davis, S. J., Foster, K. M., Jalca, A., Meyer, B. H., Mizen, P., Navarrete, M. A., Smietanka, P., Thwaites, G., and Wang, B. Z. (2026). Firm data on ai. Working Paper 34836, National Bureau of Economic Research.

Yu, Y., Fleming, M., Hampton, L., Combemale, C., and Thompson, N. (2026). Ai adoption in s&p 500 firms. *arXiv preprint arXiv:2607.08920*.

Appendix

A Data, sample construction, and variable definitions

The estimation sample begins from the full Capital IQ transcript universe restricted to fiscal years 2011 through 2024. Table A.1 traces the funnel from that universe to the firm-year estimation panel. The strict deployment-disclosure event is the first transcript scored 1 by the evidence-gated LLM classifier; the exact disclosure timestamp is aligned to Compustat fiscal periods as described in Appendix A.1.

Table A.1: Sample construction and flow

Stage	Count
Panel A. Transcript measurement pipeline	
Unique corporate transcript events dated 2011–2024	187,191
Raw keyword-screened transcript–company records	23,459
Deduplicated events with both AID and GenAID scores	22,974
AI deployment or exploration ($AID \geq 0.5$)	16,376
Strict AI deployment ($AID = 1$)	12,617
GenAI deployment or exploration ($GenAID \geq 0.5$)	3,779
Strict GenAI deployment ($GenAID = 1$)	2,630
Panel B. Estimation panel	
Firm-year observations	40,404
Unique firms	3,898
Firms with a stock-aligned strict-disclosure cohort	1,293
first movers / followers	383 / 910

Notes. Panel A distinguishes raw transcript–company extraction records from unique corporate events. The keyword screen produces 23,459 raw records before alternate transcript versions are collapsed. These records map to 22,974 unique corporate events, all of which carry both AID and GenAID scores. The score-class counts are calculated after retaining one record per corporate-event identifier. The transcript universe is restricted using calendar event dates; alignment to Compustat fiscal periods occurs after the transcript-level measurement is constructed. First movers are adopters at or below the within-industry 25th percentile of the predetermined diffusion state (Appendix B); no minimum-adopter screen is imposed in the headline definition. Panel B shows sample used in the discrete mover comparisons which contains 1,293 adopters: 383 first movers and 910 followers. Continuous-diffusion specifications use the predetermined diffusion state directly and report their own outcome-specific treated-firm support.

Valuation measures. The primary outcome is raw book-assets q. Two accounting-perimeter measures capitalize internally generated intangibles: (i) a denominator-only intangible-adjusted q, which adds the estimated replacement value of internally generated

knowledge and organization capital to the denominator only, and (ii) Peters–Taylor total q, which scales operating value by gross physical plus intangible replacement capital after subtracting current assets. All three are winsorized by fiscal year at the 1st and 99th percentiles. Where internal capital is positive, the adjusted measure is mechanically below raw q in 100% of observations (the denominator grows while the numerator is fixed); the raw–adjusted correlation is 0.789. The custom Peters–Taylor series correlates 0.976 with the official WRDS `totalq` series on the overlap.

A.1 Disclosure timing, observability, and risk-set samples

Exact disclosure timestamps are mapped to the first fiscal year-end on or after the disclosure date (the stock cohort) and to the first fiscal period beginning after disclosure (the flow cohort). Treatment is absorbing. A missing transcript period is not treated as an observed non-disclosure; the median gap between a firm's last observed non-strict transcript and its first strict disclosure is 160 days.

Table A.2 reports the observation and timing samples used for the first-mover/follower robustness analysis. The baseline uses all firm-years. The current-period-observable sample keeps only firm-years with an observed transcript in the period, while the right-censored sample drops firm-years after a firm's final observed transcript. These two restrictions address the untreated risk set. A separate timing-restricted sample retains all never-treated firms but keeps an adopter only when the absolute interval between its last observed non-strict disclosure and first strict disclosure is nonmissing and no more than 365 days. This last restriction addresses imprecision in the dating of the deployment transition rather than transcript observability. Table A.3 reports retention through event time.

Figure A.1 summarizes the observation process behind the risk-set samples. Panel A shows that transcript coverage of the firm-year panel improves over time, from about 68% in the mid-2010s to 92% by 2024, so the sample is progressively more complete. Panel B plots the cumulative distribution of the gap between a firm's last observed non-strict transcript and its first strict-deployment disclosure. The median disclosure gap is 160 days,

Table A.2: Risk-set and timing-precision samples used in the first-mover/follower estimator

Sample	Firm-years	Firms	Classified adopters
Baseline	40,404	3,898	1,293
Current-period observable	31,578	3,708	1,293
Right-censored	32,510	3,710	1,293
Disclosure gap $\leq$ 365 days	32,089	3,120	517

Notes: Each sample is passed to the untreated-outcome imputation procedure. The final column reports the classified adopters available for the first-mover/follower comparison in each sample. The current-period-observable and right-censored samples alter the observability of untreated firm-years. The final row retains all never-disclosing firms but restricts adopters to the 517 firms with a nonmissing absolute disclosure interval of no more than 365 days. The same predetermined classification rule and firm-cluster bootstrap are used throughout.

Table A.3: Retention through event time, by first-mover status

Event time k	0	1	2	3	4	5	6	7
First movers	1.000	0.990	0.981	0.971	0.970	0.975	0.971	0.963
Followers	1.000	0.981	0.964	0.951	0.966	0.959	0.957	0.952

Notes. Share of calendar-feasible adopters observed at each event horizon.

and roughly three-quarters of gaps are no more than one year, so the transition in observed deployment status is dated with moderate precision. Panel C shows that retention through event time is high and nearly identical for first movers (96.3% at $k = 7$) and followers (95.2%), so the medium-run comparison is not shaped by differential attrition.

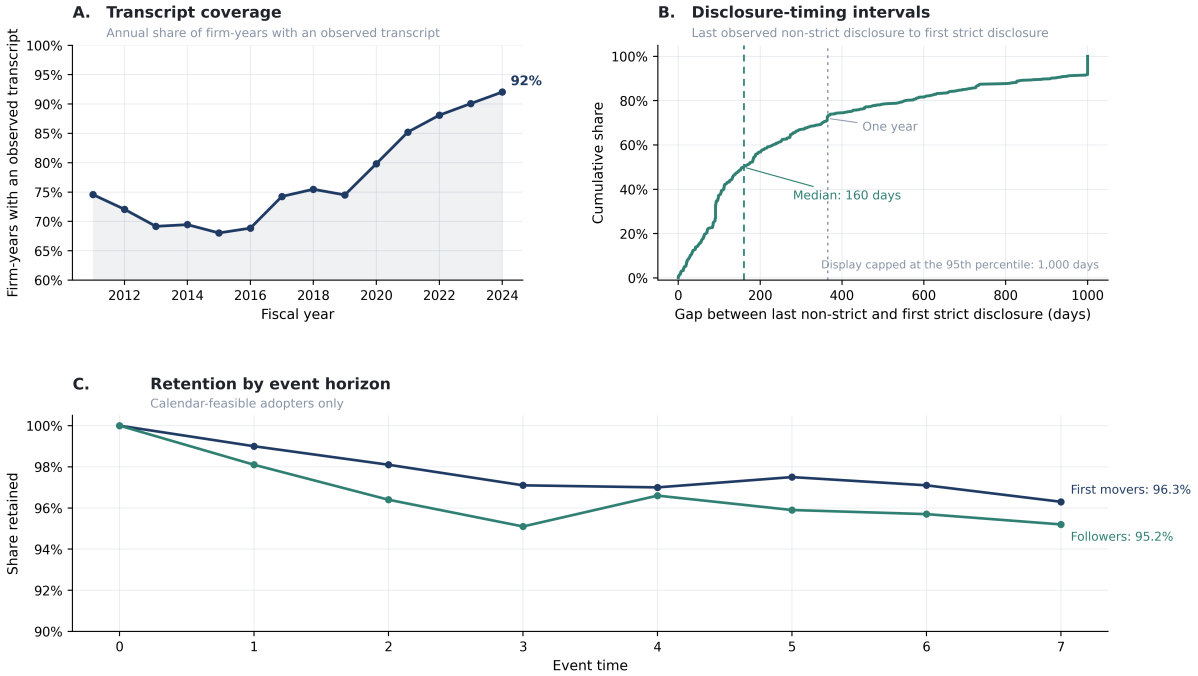


Figure A.1: **Observation coverage, disclosure timing, and event-horizon retention.** Panel A: annual share of firm-years in the estimation panel with an observed transcript. Panel B: cumulative distribution of the gap, in days, between the last observed non-strict AI disclosure and the first strict-deployment disclosure (display capped at the 95th percentile, 1,000 days; median 160 days, dashed line; one year, dotted line). Panel C: share of adopters retained by event time, separately for first movers and followers, among calendar-feasible adopters. Coverage improves over the sample, disclosure gaps are generally moderate, and retention is similar across mover groups.

B First-mover/follower balance and support

The predetermined diffusion state S_i is the lag-one, all-sample-firms leave-one-out industry adoption share read at the exact cohort year, i.e. the peer adoption share at $G_i - 1$. First movers are adopters at or below the within-industry adopter 25th percentile of S_i. Table B.1 reports covariate balance across the split. First movers are materially larger; the split is not random and is interpreted as a predetermined diffusion position.

Table B.1: First-mover/follower balance

Covariate	First movers	Followers	Normalized diff.
Size (log assets)	8.115	6.944	0.469*
q	2.672	2.958	−0.120
Profitability	−5.674	−7.092	0.015
R&D / assets	0.094	0.113	−0.130

Notes. Pre-adoption means. * marks |normalized difference| > 0.25. Adopters are spread across 415 industry-by-cohort-year cells (mean 3.12 adopters per cell; 78.8% of cells hold fewer than five adopters), which is why the continuous predetermined-diffusion design in the main text, rather than a saturated industry-by-cohort interaction, is the workhorse.

C Event-study robustness

Table C.1 reports three pre-disclosure diagnostics for each group and outcome. For the valuation outcomes, all local near-event tests and linear-drift tests fail to reject. The all-lead tests reject for followers on intangible-adjusted and Peters–Taylor q, but these rejections are generated by the pooled distant endpoint rather than the near-event leads. Among the real-side outcomes, the principal pre-period concerns arise for intangibles/assets and the follower markup path. These diagnostics motivate the paper's emphasis on the local valuation evidence and its more descriptive interpretation of the real-side event studies.

Table C.1: Pre-trend joint tests, by outcome and diffusion group

Outcome	Group	Local joint	All-lead joint	Linear drift
Panel A. Valuation (Section 5)				
Raw q	First movers	0.92	0.94	0.96
	Followers	0.37	0.49	0.46
Intangible-adjusted q	First movers	0.95	0.73	0.76
	Followers	0.21	0.01	0.90
Peters–Taylor total q	First movers	0.84	0.78	0.98
	Followers	0.19	<0.01	0.36
Panel B. Real outcomes (Sections 7, IA.15)				
Intangibles / assets	First movers	0.17	0.08	0.04
	Followers	<0.01	<0.01	<0.01
R&D / assets	First movers	0.99	0.31	0.76
	Followers	0.48	0.44	0.40
Operating margin	First movers	0.50	0.56	0.55
	Followers	0.13	0.23	0.30
Cost of goods / sales	First movers	0.24	0.37	0.62
	Followers	0.57	0.73	0.47
Production-based markup	First movers	0.35	0.45	0.41
	Followers	<0.01	<0.01	0.60

Notes: Joint tests that the pre-adoption event-time coefficients are zero, from the group-specific Sun–Abraham event studies against the common never-adopter pool. The *local joint* test uses the near-event leads and is the primary test; the *all-lead joint* test additionally includes the pooled $k \leq -5$ endpoint bin, which is compositionally distinct and can reject on distant pre-history; the *linear drift* test is a one-degree-of-freedom test for a linear pre-adoption slope. Cluster-F p-values, clustered by firm; values below 0.01 shown as "<0.01". For the valuation outcomes, the local and linear-drift tests are clean for both groups. The all-lead tests reject for followers on intangible-adjusted and Peters–Taylor q, where the rejection is driven by the pooled distant endpoint. Pre-period rejections among the real-side outcomes are concentrated in intangibles/assets and the follower markup path.

D Capability by diffusion across the Peters–Taylor grid

Table D.1 reports the cross-sectional interaction between pre-adoption capability and the predetermined diffusion state, Δq on capability A, diffusion S, and $A \times S$, with controls and industry-plus-cohort fixed effects, clustered by industry. The theory predicts a negative interaction (the value of scarce capability compresses as peers adopt). The result is negative and significant under sector fixed effects across all Peters–Taylor calibrations and both capability proxies; it attenuates and loses significance under fine-industry fixed effects, which is reported as a sensitivity rather than treated as purely mechanical. The interaction is negative across all reported Peters–Taylor parameterizations under sector fixed effects and attenuates materially under fine-industry fixed effects.

Table D.1: Capability–diffusion interaction across valuation perimeters and capitalization calibrations

Panel A. Baseline calibration across valuation perimeters					
		Sector fixed effects		Fine-industry fixed effects	
Valuation perimeter	Capability proxy	Coef.	t	Coef.	t
Raw q	K^{int}/AT	−0.282	−2.43	−0.156	−1.10
	R&D/AT	−0.206	−0.95	−0.031	−0.14
Intangible-adjusted q	K^{int}/AT	−0.166	−2.82	−0.095	−1.31
	R&D/AT	−0.155	−1.70	−0.060	−0.67
Peters–Taylor total q	K^{int}/AT	−0.386	−3.20	−0.159	−1.19
	R&D/AT	−0.583	−4.96	−0.200	−1.37
Panel B. Peters–Taylor capitalization sensitivity					
Parameterization		K^{int}/AT		R&D/AT	
		Coef.	t	Coef.	t
$\lambda = 0.20$		−0.434	−3.27	−0.658	−5.12
$\lambda = 0.40$		−0.343	−3.09	−0.512	−4.68
$\delta_O = 0.15$		−0.369	−3.15	−0.557	−4.86
$\delta_O = 0.33$		−0.418	−3.29	−0.634	−5.14
$\delta_R = 0.20$		−0.429	−3.22	−0.638	−4.95
$\delta_R = 0.30$		−0.501	−3.22	−0.731	−4.85

Notes: The table reports coefficients on the interaction between pre-disclosure capability and the predetermined diffusion state. Panel A compares valuation perimeters under the headline capitalization parameters using sector and fine-industry fixed effects. Panel B reports the Peters–Taylor capitalization sensitivity under sector fixed effects, varying the SG&A capitalization share, λ, and the depreciation rates of organization and knowledge capital, δ_O and δ_R. Standard errors are clustered by industry.

Internet Appendix

IA.1 Measurement, extraction, and validation

IA.1.1 Transcript universe: counts and event-type composition

Table IA.1 reports annual transcript counts in the full corpus. Table IA.2 summarizes event types. The corpus is dominated by earnings calls, with substantial coverage of conference presentations and shareholder/analyst calls.

Table IA.1: Annual transcript measurement-pipeline counts

Year	Full unique events	Raw screened records	Deduplicated scored events	$AID \geq 0.5$	$AID = 1$	$GenAID \geq 0.5$	$GenAID = 1$
2011	8,913	63	60	24	20	0	0
2012	10,151	87	84	24	19	1	0
2013	9,976	100	98	40	28	0	0
2014	10,219	181	175	75	58	0	0
2015	10,453	281	280	132	103	1	1
2016	10,507	651	640	320	257	10	6
2017	12,780	1,503	1,468	946	680	16	12
2018	13,157	1,950	1,882	1,277	923	21	16
2019	13,605	2,153	2,114	1,450	1,075	32	29
2020	16,706	2,536	2,484	1,514	1,176	43	37
2021	18,134	3,240	3,153	2,003	1,574	54	47
2022	17,378	2,770	2,733	1,721	1,418	72	62
2023	17,817	3,953	3,890	3,354	2,544	1,624	978
2024	17,395	3,991	3,913	3,496	2,742	1,905	1,442
Total	187,191	23,459	22,974	16,376	12,617	3,779	2,630

Notes. Full unique events count corporate transcript events dated 2011–2024. Raw screened records count transcript–company records before version deduplication. Deduplicated scored events retain one record per corporate event; all such events carry both AID and GenAID scores. The final four columns report classifications in this deduplicated event-level sample.

Table IA.2: Event-type distribution in the 2011–2024 transcript universe

Event type	Count
Earnings Calls	120,388
Company Conference Presentations	43,924
Shareholder/Analyst Calls	7,645
Special Calls	7,355
Analyst/Investor Day	3,614
M&A Calls	2,913
Sales/Trading Statement Calls	755
Guidance/Update Calls	450
Fixed Income Calls	101
Operating Results Calls	29
Interim Management Statement Calls	17
Total	187,191

Notes: Counts cover the 187,191 unique corporate transcript events from 2011–2024. Event-type categories are mutually exclusive and exhaustive.

IA.1.2 Keyword dictionary and normalization

I apply a deterministic transcript-level keyword screen to route potentially AI-relevant transcripts into structured parsing. Matching is case-insensitive and applied to raw transcript text after standard whitespace normalization. The purpose of this screen is computational: it reduces the number of transcripts sent to the structured extraction stage. Importantly, the screen is not used to define adoption. Downstream parsing assigns adoption scores and can reject residual false positives.

IA.1.2.1 Dictionary

The keyword dictionary is:

> `Artificial Intelligence`, `AI`, `Machine Learning`, `ML`, `Deep Learning`, `Neural Networks`, `Generative AI`, `Generative Models`, `Deep Generative Models`, `Chatbots`, `Virtual Assistants`, `Autonomous Systems`, `AI Integration`, `OpenAI`, `GenAI`, `ChatGPT`, `Copilot`, `language model`, `Natural Language Processing`, `autonomous vehicles`, `AI Research`, `Digital Transformation`, `Predictive Analytics`.

For short tokens such as `AI` and `ML`, matches are restricted to word-boundary occurrences (e.g., regex-style `\bAI\b`) to avoid mechanical false positives from substrings in common words or identifiers.

IA.1.3 Prompt schema and scoring examples

IA.1.3.1 Unit of analysis and target-company isolation

Extraction is performed at the (transcript_id, companyid) level. This matters because transcripts often mention multiple firms (competitors, suppliers, customers, and partners). The extraction protocol therefore conditions all questions on a target company and treats non-target firm mentions as context rather than attribution, reducing cross-firm misclassification.

IA.1.3.2 Output record and missingness conventions

Each transcript–company pair is mapped to a single machine-readable record (JSON) with invariant identifiers (`companyid`, `companyname`) and one key–value field per question. When the transcript does not contain sufficient evidence to populate a field, the output must take an explicit null value (e.g., `Information not found` or `Cannot infer from context`). This convention makes missingness explicit and supports systematic auditing.

IA.1.3.3 Question blocks

The question set has two blocks. Descriptive fields characterize the nature of AI activity (ecosystem role, sourcing, application type, business-unit context, and short summaries). Numerical fields produce scalar measures used in the main analysis, most importantly adoption scores for AI and Generative AI.

IA.1.3.4 Evidence-gated adoption scoring (extensive margin)

I map each transcript to an evidence-gated adoption-extent score—used throughout the paper under the variable name AID (AI Deployment). Formally, for firm i in period t, the score $AID_{i,t} \in \{0, 0.5, 1\}$ is assigned by the LLM under the additional requirement that any nonzero value must be accompanied by one or more supporting sentences quoted from the transcript. This evidence-gating rule makes the classification directly traceable to the underlying text, strengthening internal validity and creating an auditable coding trail.

The scoring rubric is intentionally coarse and targets the extensive margin of deployment: (i) $AID_{i,t} = 1$ is assigned only when the transcript contains clear evidence of deployment or integration (e.g., implemented in products/services or embedded in internal operations); (ii) $AID_{i,t} = 0.5$ is assigned when the transcript reflects exploration without confirmed integration (e.g., pilots, governance/process references, experimentation, vendor evaluations, or stated intentions); (iii) $AID_{i,t} = 0$ is assigned when no evidence of deployment or integration is present.

Figure IA.1 summarizes the implied adoption extent over time within the LLM-coded (keyword-filtered) transcript sample. Specifically, the stacked bars plot, by year, the

share of firm-years in the coded sample assigned to each score. Because transcripts enter the coded sample only if they contain AI-related keywords, the plotted shares should be interpreted as conditional distributions among AI-mentioning firm-years rather than adoption rates in the full transcript universe. Panel A reports general AI adoption extent; Panel B reports GenAI adoption extent.

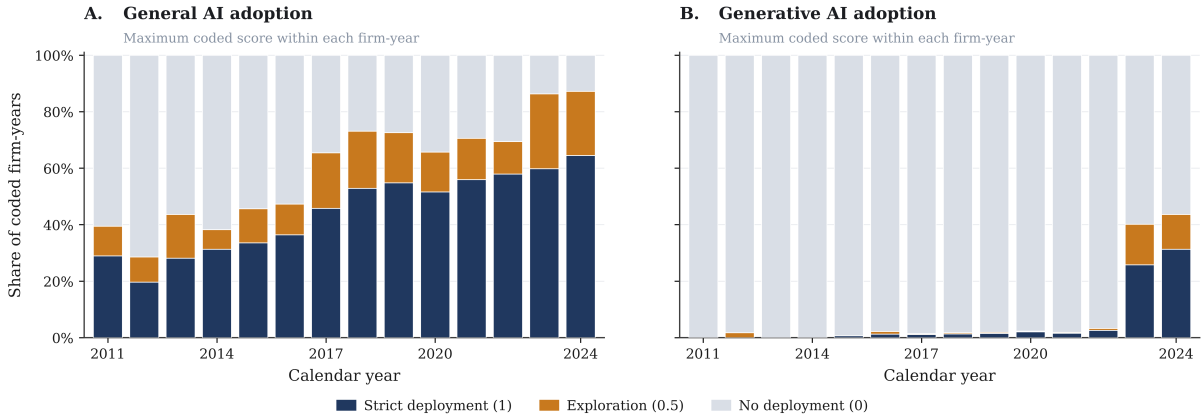


Figure IA.1: Adoption extent implied by evidence-gated scoring (AID), conditional on the LLM-coded sample. Stacked bars show, by year, the share of firm-years assigned $AID \in \{0, 0.5, 1\}$, where nonzero values require one or more transcript sentences as supporting evidence. Panel A shows general AI adoption extent; Panel B shows GenAI adoption extent.

IA.1.3.5 Execution and parsing

Extraction is executed in batches with checkpointing and periodic writes to disk. Outputs are parsed by extracting the top-level JSON object and validating field presence and admissible null values prior to aggregation.

IA.1.3.6 Coverage distribution

Over the broad coverage panel of 4,313 firms across 2011–2024 (60,382 firm–years), 33,460 firm–years, or 55.4%, contain at least one transcript. This pooled share is deflated by inactive and pre-entry cells: it counts every cell in the balanced firm-by-year grid, including years before a firm enters the sample or after it exits, and years in which an active firm simply held no call. Within the estimation panel used in the event studies, per-period coverage is considerably higher, between about 68% and 92% by year and rising to 92% by 2024 (main text, Section 2.4). Figure IA.2 shows the distribution of transcripts per

firm–year in the merged analysis dataset.

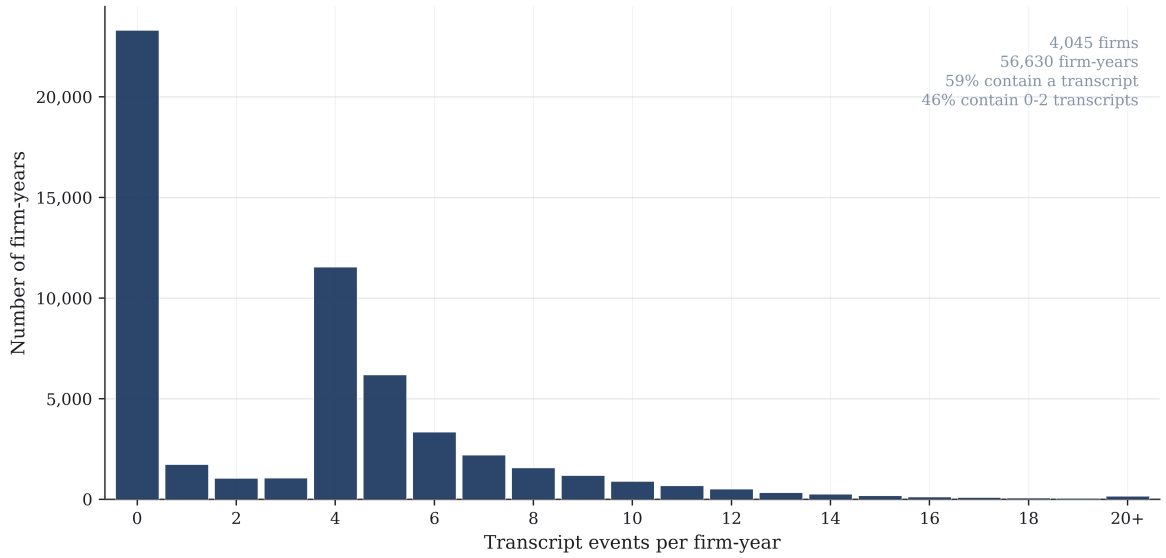


Figure IA.2: **Distribution of transcript coverage across firm-years.** The histogram reports transcript-event counts in the broad 2011–2024 firm-by-year coverage panel of 4,313 firms and 60,382 firm-years, of which 33,460 contain at least one transcript. Corporate events are deduplicated within firm and event identifier. Counts are capped at 20, and the final bar contains all firm-years with 20 or more transcript events. The mass at zero partly reflects inactive and pre-entry cells in the broad coverage grid.

IA.1.3.7 Illustrative scoring cases

To make the evidence-gated scoring logic transparent, I report four representative examples of transcript evidence alongside the corresponding structured outputs. The examples span the full adoption scale ($AID \in \{0, 0.5, 1\}$) and different ecosystem roles (internal user vs. vendor). For each case, I show the relevant evidence text and the key output fields used in the analysis (with nonzero AID supported by one or more transcript sentences, per the evidence-gating rule).

Illustration I: No adoption vs. exploratory GenAI. Figure IA.3 contrasts (i) AI discussed as an external demand driver ($AID = 0$) with (ii) exploratory GenAI pilots and plans without demonstrated deployment or integration ($AID = 0.5$). The key distinction is that selling into AI-driven demand ("picks-and-shovels") does not constitute adoption absent evidence of internal deployment or product/operational integration.

Score 0 Example: Micron Technology, Inc.
Narrative (paraphrased):

AI and autonomous vehicles are major demand drivers for DRAM and NAND. Customers' AI workloads require more memory and storage.

LLM-derived labels (selected):

- **AI adoption score:** 0.
- **AI description:** AI discussed as a *demand driver* for memory products, not an internal capability.
- **AI Application Type (Inferred):** *Cannot infer from context*
- **AI Strategy:** *Cannot infer from context*

Score 0.5 Example (GenAI): Coursera, Inc.
Transcript evidence (paraphrased):

Coursera plans to use generative AI and models like ChatGPT to create interactive, personalized learning experiences and hopes to launch something "pretty special" for learners.

LLM-derived labels:

- **AI/Gen AI adoption score:** 0.5.
- **Application types:** Generative & Creative AI, Interactive AI.
- **Technologies:** Generative AI, language models.
- **Role:** Mixed user/vendor; AI features embedded in the learning platform and leveraging external models (`AI_Source = others`).

Figure IA.3: Illustration I: No adoption vs. exploratory GenAI (paraphrased transcript evidence). The left panel shows AI discussed solely as an external demand driver ($AID = 0$), while the right panel shows exploratory GenAI plans and pilots without demonstrated deployment or integration ($AID = 0.5$).

Illustration II: Strong internal user vs. vendor. Figure IA.4 illustrates why role and sourcing fields complement the adoption score. Both cases exhibit confirmed deployment/integration ($AID = 1$), but differ in ecosystem role. The Coterra example is a strong internal-user case in which machine learning is described as deployed in operations. The C3.ai example is a vendor case: the firm develops and deploys AI capabilities in products/platforms sold to customers (i.e., adoption via product integration), which is distinct from merely benefiting from AI-related demand.

Strong Internal User: Coterra Energy Inc.

Event: Barclays CEO Energy–Power Conference 2023
Transcript ID: 2903332

Transcript evidence:

"Machine learning is really democratized at Coterra. It's a part of how we work across the company and is outperforming our best reservoir engineers."

Structured LLM output (selected):

- **AI adoption score:** 1.
- **Application type:** Analytical & Predictive AI.
- **Technologies:** Machine learning models.
- **Business units:** Operations (production, project planning, well performance).
- **Role: User only** (`AI_Dev_For_Sale = 0`, `AI_Source = in-house`).

Pure Vendor: C3.ai, Inc.

Event: Analyst/Investor Day
Transcript ID: 2842374

Transcript evidence (paraphrased):

C3.ai develops and sells enterprise AI applications and a platform that sit on top of customers' data, with dozens of turnkey apps deployed across industries.

Structured LLM output (selected):

- **AI adoption score:** 1.
- **AI_Dev_For_Sale:** 1 (pure AI vendor).
- **Application types:** Analytical & Predictive, Generative, Interactive AI.
- **Technologies:** ML, deep learning, NLP, generative AI, LLMs.
- **AI source:** Developed in-house and sold as a platform to customers.

Figure IA.4: Illustration II: Strong internal user vs. vendor (paraphrased transcript evidence where noted). Both cases exhibit deployment/integration ($AID = 1$), but differ in ecosystem role: an internal operational user (left) versus a firm integrating AI into products/platforms sold to customers (right).

IA.1.4 Pipeline reliability and validation

For each transcript–firm observation, the pipeline outputs (i) an evidence-gated adoption/deployment extent score for general AI, denoted AID and taking values in $\{0, 0.5, 1\}$, and (ii) a supporting excerpt intended to justify any nonzero classification. I construct an analogous GenAI score, denoted GenAID $\in \{0, 0.5, 1\}$, using the same rubric and evidence-gating rule. The design is auditable: by construction, nonzero scores must be accompanied by text support, and the zero class returns an empty excerpt.

Evidence availability. Table IA.3 summarizes supporting-evidence availability across the 23,459 raw transcript–company extraction records. The unit is one LLM extraction call before alternate transcript versions are collapsed to unique corporate events. This table is therefore an implementation diagnostic rather than an analysis-sample count.

Table IA.3: Supporting evidence availability by score class in the raw parsed-record sample

Indicator	Score class	N	N with support	Share with support
AI Adoption	0.0	6,828	0	0.000
AI Adoption	0.5	3,810	3,810	1.000
AI Adoption	1.0	12,821	12,821	1.000
GenAI Adoption	0.0	19,634	0	0.000
GenAI Adoption	0.5	1,162	1,162	1.000
GenAI Adoption	1.0	2,663	2,663	1.000

Notes. Counts refer to the 23,459 raw transcript–company records submitted to the LLM, before corporate-event deduplication. Nonzero scores require supporting evidence; the zero class returns no excerpt by design.

Automated anchoring of supporting excerpts (evidence locatability). A key failure mode in LLM pipelines is fabricated supporting text. My prompts instruct the model to return an excerpt drawn from the transcript (allowing minor formatting differences), and I verify at scale whether each excerpt can be located in the raw transcript text. Because excerpts may differ in punctuation, capitalization, or may concatenate adjacent sentences, I use a hierarchical matching procedure. First, I attempt normalized substring matches after case-folding and whitespace normalization. For remaining cases, I apply a conservative windowed fuzzy match over short transcript windows (up to three consecutive sentences). Excerpt–transcript pairs that do not meet these automated criteria are flagged for manual inspection.

Table IA.4 reports anchoring rates under a strict definition and a more permissive locatability definition. Under the strict definition, 97.49% of supported excerpts are anchored; allowing moderate-threshold fuzzy matches raises locatability to 99.08%.

Table IA.4: Transcript anchoring rates for supporting excerpts

Anchoring definition	Share anchored
Strict anchor (substring or high-threshold windowed fuzzy)	0.9749
Overall locatability (strict anchor or fuzzy ≥ 70)	0.9908

Notes: The strict definition requires either a normalized substring match or a high-threshold windowed fuzzy match over up to three consecutive transcript sentences. "Overall locatability" additionally counts moderate-threshold fuzzy matches (score ≥ 70) as locatable.

Placebo abstention checks (falsification / fabrication propensity). As a falsification diagnostic, I include three placebo items designed to elicit abstention unless the transcript contains explicit support. For each item, the instructed response when no evidence is present is the string `Information not found` (minor formatting variants are treated as equivalent). Placebo items do not enter the adoption indicators; they are used only to diagnose fabrication when abstention is required:

- **Fictional AI use case:** references to a non-existent system (e.g., `X-QuantumMind`).
- **Secret AI department:** claims of a stealth or undisclosed AI division.
- **Quantum AI investment:** discussion of "quantum AI" spending (not a strict placebo because transcripts occasionally mention quantum-related initiatives).

Table IA.5 reports abstention rates across all 23,459 transcript records processed by the classification pipeline. Some records correspond to the same corporate event, so this diagnostic is calculated before the data are collapsed to 22,974 unique corporate events. Abstention is near-universal for the two strict placebos. The "Quantum AI Investment" item exhibits slightly lower abstention, consistent with occasional genuine mentions of quantum-related topics.

Table IA.5: Placebo abstention rates

Placebo item	N	Share `Information not found`	Share non-empty (non-notfound)
Fictional AI Use Case	23,459	0.9955	0.0024
Secret AI Department	23,459	0.9962	0.0017
Quantum AI Investment (not a strict placebo)	23,459	0.9928	0.0051

Notes: `Information not found` is treated as abstention, with minor formatting variants (e.g., trailing punctuation) collapsed. "Share non-empty" excludes empty strings; the residual share corresponds to empty/missing outputs. "Quantum AI Investment" is not a strict placebo because transcripts occasionally discuss quantum computing or related initiatives; its non-abstentions may reflect genuine mentions rather than fabrication.

Manual audit and validation against human coding. I manually audited two rounds of transcript–firm observations (total $N = 455$) using the same rubric as the automated pipeline. The audit had two objectives: (i) target attribution (verifying that the supporting excerpt refers to the focal firm (rather than third parties mentioned in the transcript)) and (ii) score correctness (assigning discrete scores in $\{0, 0.5, 1\}$ for AI Adoption (AID) and GenAI Adoption (GenAID)). I treat the human label as the reference

and compare it to the LLM score. Tables IA.6 and IA.7 report confusion matrices (LLM score in rows; human score in columns), along with class-wise precision, recall, and F1.

Table IA.6: Manual validation of AI Adoption score against human coding

	Human score			Class-wise performance (human as reference)		
LLM score	0.0	0.5	1.0	Precision	Recall	F1
0.0	61	8	7	0.803	0.938	0.865
0.5	4	56	18	0.718	0.848	0.778
1.0	0	2	299	0.993	0.924	0.957
Overall accuracy (3-class)					0.914	
Adjacent disagreement rate (0↔0.5 or 0.5↔1)					0.070	
Non-adjacent disagreement rate (0↔1)					0.015	
0→1 mistake rate (LLM=1, Human=0)					0.000	
Binary adoption (score ≥ 0.5)				0.989	0.962	0.975

Notes: The table combines two audit rounds for a total of $N = 455$ transcript–firm observations. Rows denote the LLM score and columns denote the human score. Adjacent disagreements treat the scale as ordinal. Binary adoption collapses scores to adopters (score ≥ 0.5) vs. non-adopters (0.0).

For AI Adoption, three-class agreement is high: overall accuracy is 0.914 (Table IA.6). Disagreements are predominantly adjacent on the ordinal scale (0 vs. 0.5 or 0.5 vs. 1): the adjacent-disagreement rate is 0.070, while non-adjacent 0-vs-1 reversals occur in only 0.015 of observations. The score-1 category is highly precise (precision 0.993) with strong recall (0.924), implying that confirmed adopters are rarely false positives. As expected, the intermediate score-0.5 category concentrates ambiguity (precision 0.718, recall 0.848). Collapsing to a binary adoption indicator (score ≥ 0.5) yields precision 0.989 and recall 0.962.

For GenAI Adoption, agreement is higher: overall three-class accuracy is 0.967 (Table IA.7). Adjacent disagreements are rare (rate 0.024) and non-adjacent 0-vs-1 reversals occur in 0.009 of cases. The score-1 category has precision 0.903 and recall 0.949, while the intermediate score-0.5 category again concentrates ambiguity (precision 0.720, recall 0.818). Under a binary GenAI adoption definition (score ≥ 0.5), precision is 0.908 and recall is 0.975.

Table IA.7: Manual validation of GenAI Adoption score against human coding

	Human score			Class-wise performance (human as reference)		
LLM score	0.0	0.5	1.0	Precision	Recall	F1
0.0	366	2	0	0.995	0.979	0.986
0.5	4	18	3	0.720	0.818	0.766
1.0	4	2	56	0.903	0.949	0.926
Overall accuracy (3-class)					0.967	
Adjacent disagreement rate (0↔0.5 or 0.5↔1)					0.024	
Non-adjacent disagreement rate (0↔1)					0.009	
0→1 mistake rate (LLM=1, Human=0)					0.009	
Binary adoption (score ≥ 0.5)				0.908	0.975	0.940

Notes: The table combines two audit rounds for a total of $N = 455$ transcript–firm observations. Rows denote the LLM score and columns denote the human score. Adjacent disagreements treat the scale as ordinal. Binary adoption collapses scores to adopters (score ≥ 0.5) vs. non-adopters (0.0).

IA.2 Outcome construction details

IA.2.1 Intangible capital

I construct internally generated knowledge and organization capital from expensed investment following Peters and Taylor (2017). Let XRD_{it} denote R&D expenditure and XSGA_{it} selling, general, and administrative expenditure. Because Compustat may include R&D in SG&A, I define cleaned SG&A as

$$\widetilde{\text{XSGA}}_{it} = \begin{cases} \max\{\text{XSGA}_{it} - \text{XRD}_{it}, 0\}, & \text{when both items are separately reported,} \\ \text{XSGA}_{it}, & \text{otherwise.} \end{cases}$$

Knowledge investment equals XRD_{it}, while organization investment equals $\lambda\widetilde{\text{XSGA}}_{it}$. The baseline organizational investment share is $\lambda = 0.30$.

Perpetual-inventory stocks. Knowledge and organization capital evolve according to

$$K^R_{it} = (1 - \delta_R)K^R_{i,t-1} + \text{XRD}_{it}, \tag{IA.1}$$

$$K^O_{it} = (1 - \delta_O)K^O_{i,t-1} + \lambda\widetilde{\text{XSGA}}_{it}. \tag{IA.2}$$

For a firm's first usable observation t_0, I initialize the two stocks separately:

$$K^R_{i,t_0} = \frac{\mathrm{XRD}_{i,t_0}}{g + \delta_R}, \tag{IA.3}$$

$$K^O_{i,t_0} = \frac{\lambda \widetilde{\mathrm{XSGA}}_{i,t_0}}{g + \delta_O}, \tag{IA.4}$$

where g is the common pre-sample investment-growth rate. I use BEA industry-specific R&D depreciation rates when available and $\delta_R = 0.15$ otherwise. The baseline organization-capital depreciation rate is $\delta_O = 0.20$.

The internally generated intangible-capital stock is

$$K^{\mathrm{int}}_{it} = K^R_{it} + K^O_{it}. \tag{IA.5}$$

To construct the broader stock used in Peters–Taylor total q, I add purchased and acquired intangible assets already reported on the balance sheet:

$$K^{\mathrm{tot}}_{it} = K^{\mathrm{int}}_{it} + \max\{\mathrm{INTAN}_{it}, 0\}. \tag{IA.6}$$

I retain K^{int}_{it} and reported intangible assets separately throughout. The denominator-only adjustment below adds only internally generated capital to book assets, avoiding double counting intangibles already included in AT_{it}. Alternative values of λ, δ_R, and δ_O are used in the capitalization and capability sensitivity analyses.

IA.2.2 Valuation-measure implementation

The three valuation ratios used in the analysis are defined in Equations (19)–(21). This subsection documents the accounting implementation underlying those definitions.

Raw book-assets q uses

$$\mathrm{AT}_{it} + \mathrm{MVE}_{it} - \mathrm{CEQ}_{it} - \mathrm{TXDB}_{it}$$

as the book-based market value of assets, with missing TXDB_{it} set to zero. The same

construction is used throughout the paper.

The denominator-only intangible adjustment adds the internally generated stock K_{it}^{int} constructed in Section IA.2.1 to book assets while holding the numerator fixed. This treatment follows the accounting logic that capitalizing previously expensed internally generated investment raises measured assets and book equity by the same amount, leaving the book-based market-value numerator unchanged. The resulting ratio is designed specifically to test whether omitted internally generated capital mechanically accounts for the raw-q differential. It is not a direct estimate of rents or of the market value assigned to each unit of intangible capital.

Peters–Taylor total q uses a broader operating-capital perimeter. Its numerator is operating firm value, $\text{MVE}_{it} + \text{DEBT}_{it} - \text{ACT}_{it}$, and its denominator is gross physical capital plus total intangible replacement capital, $\text{PPEGT}_{it} + K_{it}^{\text{tot}}$. Here, K_{it}^{tot} combines internally generated capital with purchased and acquired intangible assets already reported in INTAN_{it}. The reconstructed series correlates 0.976 with the official WRDS `totalq` series on their overlapping sample, which provides an external reconciliation of the implementation.

All three ratios are winsorized within fiscal year at the first and ninety-ninth percentiles.

IA.2.3 Realized operating outcomes

The realization analysis uses a pre-specified family of six operating outcomes. Define raw operating profit as

$$\pi_{it}^{\text{raw}} = \text{OIADP}_{it}. \tag{IA.7}$$

Because OIADP_{it} is measured after expensing R&D and organizational investment, I construct intangible-adjusted operating profit by adding current intangible investment and subtracting economic depreciation on beginning-of-period capital:

$$\pi_{it}^{\text{adj}} = \text{OIADP}_{it} + \text{XRD}_{it} + \lambda\widetilde{\text{XSGA}}_{it} - \delta_R K_{i,t-1}^{R} - \delta_O K_{i,t-1}^{O}. \tag{IA.8}$$

The headline realization outcomes use $(\lambda, \delta_R, \delta_O) = (0.30, 0.15, 0.20)$, with the applicable BEA R&D depreciation rate replacing 0.15 when available.

Net operating assets are

$$\mathrm{NOA}_{it} = \mathrm{AT}_{it} - \mathrm{CHE}_{it} - \big(\mathrm{LCT}_{it} - \mathrm{DLC}_{it}\big). \tag{IA.9}$$

I measure invested capital using beginning- and end-of-period averages:

$$\overline{\mathrm{NOA}}_{it} = \frac{\mathrm{NOA}_{it} + \mathrm{NOA}_{i,t-1}}{2}, \tag{IA.10}$$

$$\overline{K}^{\mathrm{int}}_{it} = \frac{K^{\mathrm{int}}_{it} + K^{\mathrm{int}}_{i,t-1}}{2}. \tag{IA.11}$$

Using a statutory tax rate $\tau = 0.21$, the six outcomes are:

(i) raw operating margin,

$$\frac{\pi^{\mathrm{raw}}_{it}}{\mathrm{Sales}_{it}};$$

(ii) intangible-adjusted operating margin,

$$\frac{\pi^{\mathrm{adj}}_{it}}{\mathrm{Sales}_{it}};$$

(iii) operating cash flow over assets,

$$\frac{\mathrm{OANCF}_{it}}{\mathrm{AT}_{it}};$$

(iv) raw return on invested capital,

$$\frac{(1-\tau)\pi^{\mathrm{raw}}_{it}}{\overline{\mathrm{NOA}}_{it}};$$

(v) intangible-adjusted return on invested capital,

$$\frac{(1-\tau)\pi^{\mathrm{adj}}_{it}}{\overline{\mathrm{NOA}}_{it} + \overline{K}^{\mathrm{int}}_{it}};$$

(vi) labor productivity,

$$\log\left(\frac{\mathrm{RealSales}_{it}}{\mathrm{EMP}_{it}}\right),$$

where RealSales_{it} is the real-sales series defined in the variable-construction appendix.

Each outcome is winsorized within fiscal year at the first and ninety-ninth percentiles. The first-mover–follower estimand, the firm-level bootstrap, and the unbalanced, balanced, and inverse-probability-of-censoring-weighted implementations are described with the realization analysis.

IA.2.4 Labor-share measures

The labor-share event studies use three firm–year ratios. I first construct a micro-based firm-level wage bill as

$$W_{it}^{\text{Firm}} = \text{Users}_{it} \times \widetilde{w}_{it},$$

where Users_{it} is firm–year headcount and $\widetilde{w}_{it}$ is median annual salary. The resulting wage bill is converted to the same monetary units as the Compustat variables.

I also construct an accounting-based labor-expense measure,

$$W_{it}^{\text{Alt}} = \begin{cases} \text{XLR}_{it}, & \text{if reported labor expense is available,} \\ 0.70 \times \text{XSGA}_{it}, & \text{otherwise.} \end{cases}$$

For Measures A and B, adjusted value added is

$$VA_{it}^{A} = \text{Sales}_{it} - \text{COGS}_{it} - 0.30 \times \text{XSGA}_{it},$$

defined only when $VA_{it}^{A} > 0$. The three labor-share outcomes are

$$\text{LS}_{it}^{A} = \frac{W_{it}^{\text{Firm}}}{VA_{it}^{A}}, \tag{IA.12}$$

$$\text{LS}_{it}^{B} = \frac{W_{it}^{\text{Alt}}}{VA_{it}^{A}}, \tag{IA.13}$$

$$\text{LS}_{it}^{C} = \frac{W_{it}^{\text{Alt}}}{\text{OIBDP}_{it} + W_{it}^{\text{Alt}}}. \tag{IA.14}$$

Measure A uses the micro-based firm-level wage bill and is available only when both

headcount and median salary are observed. Measure B uses reported labor expense when available and an SG&A-based proxy otherwise, while retaining the same adjusted value-added denominator as Measure A. Measure C uses the same reported-or-imputed labor-expense numerator as Measure B but scales it by operating income before depreciation plus labor expense. Measures A and B therefore share a denominator, while Measures B and C share a numerator; the three constructions are related robustness measures rather than independent outcomes.

IA.2.5 Markups and price–cost margins

Gross margin. I measure the accounting price–cost margin as

$$\mathrm{PCM}_{it} = \frac{\mathrm{Sales}_{it} - \mathrm{COGS}_{it}}{\mathrm{Sales}_{it}}. \tag{IA.15}$$

This measure equals reported gross profit divided by sales.

Production-based markups. I also construct a production-based markup measure that treats COGS as the flexible input in a Cobb–Douglas revenue production function. For each industry g, pooling all available years, I estimate

$$\log \mathrm{Sales}^{\$}_{it} = \beta_{K,g} \log K_{it} + \beta_{L,g} \log L_{it} + \theta_g \log \mathrm{COGS}^{\$}_{it} + \alpha_i + u_{it}, \tag{IA.16}$$

where K_{it} and L_{it} are the physical-capital and labor inputs defined in the variable-construction appendix, α_i is a firm fixed effect, and θ_g is the estimated revenue elasticity with respect to COGS.

I estimate Equation (IA.16) separately within each industry by absorbing the firm fixed effects. Industries with fewer than 100 usable firm–year observations are excluded from the production-function estimation.

Let

$$s^{\mathrm{COGS}}_{it} = \frac{\mathrm{COGS}^{\$}_{it}}{\mathrm{Sales}^{\$}_{it}} \tag{IA.17}$$

denote the COGS revenue share. The implied production-based markup is

$$\mu_{it} = \frac{\theta_g}{s_{it}^{\text{COGS}}}. \tag{IA.18}$$

Markup levels are winsorized within fiscal year at the first and ninety-ninth percentiles before aggregation or regression analysis. Log-markup specifications require $\mu_{it} > 0$. When presenting level and log event studies together, I use the same positive-markup sample to hold sample composition fixed.

Finance and Insurance variants. Interest expense can be an important variable cost for firms in NAICS 52. I therefore construct an adjusted flexible-input share:

$$s_{it}^{\text{FI}} = \begin{cases} \dfrac{\text{COGS}_{it}^{\$} + \text{XINT}_{it}^{\$}}{\text{Sales}_{it}^{\$}}, & \text{if firm } i \text{ is in NAICS 52 and } \text{XINT}_{it}^{\$} \text{ is observed,} \\ \dfrac{\text{COGS}_{it}^{\$}}{\text{Sales}_{it}^{\$}}, & \text{otherwise.} \end{cases} \tag{IA.19}$$

The corresponding adjusted markup is

$$\mu_{it}^{\text{FI}} = \frac{\theta_g}{s_{it}^{\text{FI}}}. \tag{IA.20}$$

I additionally report a specification that excludes NAICS 52 firms entirely.

IA.3 Who discloses AI deployment? Selection, prediction, and diffusion position

This appendix characterizes selection along two margins. First, I examine which firms subsequently make a first strict public disclosure of AI deployment. Second, among firms that eventually disclose, I compare the predetermined characteristics of firms disclosing at low and high positions on their industries' diffusion curves. The first exercise combines conditional-association estimates with out-of-sample prediction; the second describes selection into the paper's first-mover–follower classification. These exercises characterize

the disclosure process but do not enter the estimation of the paper's main valuation effects.

IA.3.1 Selection into first strict disclosure

Association and prediction. I separate conditional association from out-of-sample prediction. For prediction, I estimate L_2-penalized logistic models. Model performance is evaluated using nested cross-validation that assigns all observations from a firm to the same fold. All preprocessing—including imputation, standardization, fixed-effect encoding, and penalty selection—is performed within the training folds. I also report a fixed calendar holdout that estimates the model using observations through 2020 and evaluates it during 2021–2024.

For economic magnitudes, I estimate a separate unpenalized logistic association model with calendar-year and six-digit GICS-industry fixed effects and standard errors clustered by firm. The reported odds ratios therefore summarize conditional associations and should not be interpreted as coefficients from the penalized prediction model.

Discrete-time first-disclosure hazard. Let G_i denote firm i's first stock-aligned fiscal cohort associated with a strict deployment disclosure, defined by $AID = 1$. Define

$$\text{FirstDisclosure}_{it} = \mathbf{1}\{t = G_i\}.$$

Let $\mathcal{R}_{it} = 1$ identify a transcript-observable firm–year in which firm i has not previously made a strict deployment disclosure. These are the firm–years in which the transcript data provide an opportunity to observe a new strict disclosure. I estimate

$$\Pr(\text{FirstDisclosure}_{it} = 1 \mid \mathcal{R}_{it} = 1) = \Lambda\big(\lambda_t + \gamma_{g(i)} + \boldsymbol{\beta}' X_{i,t-1}\big)\,, \tag{IA.21}$$

where $\Lambda(\cdot)$ is the logistic distribution function and λ_t and $\gamma_{g(i)}$ are calendar-year and six-digit GICS-industry fixed effects.

The predictors in $X_{i,t-1}$ are measured in the preceding consecutive fiscal year and include log market capitalization, R&D/assets, raw book-assets q, intangibles/assets,

debt/assets, capital expenditures/assets, cash/assets, operating margin, disclosure intensity, and the lagged leave-one-out industry deployment share. After requiring the lagged predictors, the estimation sample contains 4,036 transcript-observable firm–years from 2,055 firms and 1,253 first-disclosure events.

Table IA.8 reports odds ratios for standardized continuous covariates. The lagged diffusion state is the strongest correlate of subsequent disclosure: a one-standard-deviation increase is associated with 2.40 times the odds of first strict disclosure. Cash/assets, intangibles/assets, and firm size are also positively associated with subsequent disclosure. The remaining accounting predictors are not statistically distinguishable from zero.

The model containing only calendar-year and industry effects achieves a firm-grouped cross-validated AUC of 0.626. Adding lagged fundamentals and the diffusion state raises the AUC to 0.656. The corresponding calendar-holdout AUC is 0.653, indicating that predictive performance changes little when the model is estimated on the earlier period and evaluated entirely during 2021–2024.

Table IA.8: Predictors of first strict AI deployment disclosure

	Odds ratio	**95% CI**	***p*-value**
Lagged industry AI diffusion	2.399	[1.892, 3.042]	< 0.001
Cash / assets	1.275	[1.129, 1.439]	< 0.001
Intangibles / assets	1.164	[1.051, 1.288]	0.003
Log market capitalization	1.132	[1.044, 1.227]	0.003
Operating margin	0.933	[0.848, 1.028]	0.162
R&D / assets	0.959	[0.872, 1.054]	0.386
Disclosure score	0.973	[0.897, 1.055]	0.506
Capital expenditures / assets	1.027	[0.941, 1.120]	0.555
Debt / assets	0.979	[0.901, 1.063]	0.613
Raw book-assets q	1.017	[0.941, 1.100]	0.665

Notes: The table reports odds ratios from an unpenalized logistic association model. Odds ratios correspond to one-standard-deviation changes in the lagged continuous covariates. The model includes calendar-year and six-digit GICS-industry fixed effects, and standard errors are clustered by firm. The usable transcript-observable sample contains 4,036 firm–years, 2,055 firms, and 1,253 first-disclosure events. These association estimates are separate from the penalized models used to evaluate out-of-sample prediction.

IA.3.1.1 Common-landmark prediction of subsequent disclosure

As a separate prospective exercise, I construct a common information set in 2018 and predict whether each firm makes its first strict deployment disclosure during 2019–2024.

Using a common landmark prevents differences in firms' prediction dates from contributing to the results. The sample contains 2,532 firms, of which 620 subsequently disclose deployment.

An industry-only ridge-logit model achieves a firm-grouped cross-validated AUC of 0.700. Adding firm fundamentals measured at the common 2018 information date raises the AUC to 0.739. Observable firm characteristics therefore contain incremental information about subsequent disclosure, although the model's overall discrimination remains moderate.

IA.3.2 Predetermined characteristics of first movers and followers

I next examine selection into the predetermined diffusion classification. For each classified discloser, I define the pre-disclosure value of a characteristic as its median over fiscal years

$$t \in \{G_i - 3, G_i - 2, G_i - 1\}.$$

The frozen classification contains 383 first movers and 910 followers. Outcome-specific samples are smaller when a characteristic is unavailable throughout the three-year pre-disclosure window.

For each standardized characteristic A_i, I estimate

$$A_i = \alpha + \beta \text{ FirstMover}_i + \gamma_{g(i)} + \varepsilon_i, \tag{IA.22}$$

where $\gamma_{g(i)}$ are six-digit GICS-industry fixed effects and standard errors are clustered by industry. The coefficient β measures the standardized first-mover-minus-follower difference conditional on industry. A second specification additionally controls for pre-disclosure log market capitalization.

I do not include disclosure-cohort fixed effects in the baseline comparison. The descriptive object is which firms disclose at lower positions on their own industries' diffusion curves, rather than which firms differ after conditioning on making their first disclosure in the same calendar cohort.

First movers enter disclosure with a larger accumulated stock of constructed intangible capital. The industry-adjusted difference in log constructed intangible capital is 0.274 standard deviations (95% CI $[0.146, 0.402]$; Holm-adjusted $p = 0.0002$). After controlling additionally for pre-disclosure firm size, the difference remains 0.209 standard deviations (95% CI $[0.082, 0.335]$; Holm-adjusted $p = 0.007$). Capital expenditures/assets are also 0.186 standard deviations higher in the industry- and size-adjusted specification (95% CI $[0.058, 0.315]$; Holm-adjusted $p = 0.022$). The Holm adjustment is applied across the family of characteristics reported in Figure IA.5 within each specification.

First movers do not have significantly higher R&D/assets, booked intangibles/assets, or constructed intangible capital/assets. The main predetermined difference is therefore in the accumulated level of constructed intangible capital, rather than in contemporaneous R&D intensity or the share of measured assets represented by intangibles.

Figure IA.5 reports the complete set of pre-disclosure comparisons.

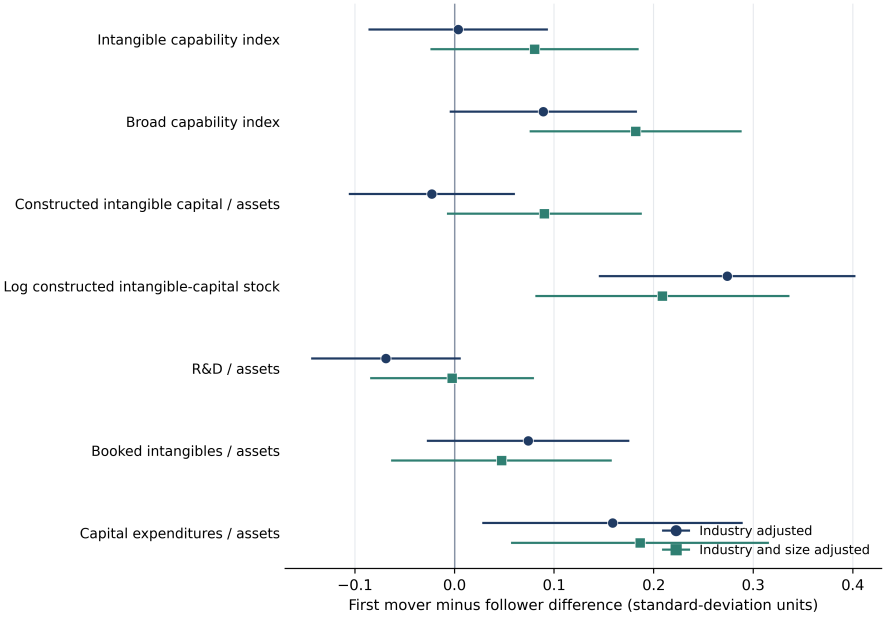


Figure IA.5: **Predetermined characteristics of first movers and followers.** The figure reports standardized first-mover-minus-follower differences in firm characteristics measured before the first strict deployment disclosure. Firm-level characteristics are medians over fiscal years $G_i - 3$ through $G_i - 1$. Circles report estimates controlling for six-digit GICS-industry fixed effects; squares additionally control for pre-disclosure log market capitalization. Horizontal bars are 95% confidence intervals based on standard errors clustered by industry. The frozen classification contains 383 first movers and 910 followers; outcome-specific estimation samples vary because of missing pre-disclosure accounting data. The baseline comparisons do not include disclosure-cohort fixed effects.

IA.4 Industry composition and diffusion

This section documents the cross-industry distribution and diffusion of strict AI-deployment disclosures. I first report the industry composition of firms with at least one strict disclosure. I then describe the cumulative strict-disclosure stock by industry for general AI and generative AI and compare diffusion paths across levels of GICS aggregation. These figures provide descriptive context for the predetermined industry diffusion state used in the main analysis.

IA.4.1 Industry composition of strict adopters

Figure IA.6 reports the industry composition of firms with at least one strict AI-deployment disclosure. Software & Services contains the largest share of strict adopters, followed by Health Care Equipment & Services and Capital Goods. At the same time, strict adopters appear across consumer, financial, industrial, health-care, and information-related industries, indicating that the measured diffusion is not confined to firms whose primary business is software production.

IA.4.2 Strict AI deployment disclosures by industry and year

Figure IA.7 reports the cumulative strict-disclosure stock within each four-digit GICS industry group. For each firm, strict-disclosure status switches on in the year of its first event with $AID = 1$ and remains on thereafter. Each industry-year cell reports the share of firms observed in that industry-year whose first strict disclosure has occurred by that year. Because the set of observed panel firms can change over time, the aggregate industry share need not be mechanically monotone even though firm-level disclosure status is absorbing.

Strict disclosures accelerate across a broad set of industries during the second half of the sample, but their timing and 2024 penetration remain heterogeneous. Software & Services has the highest 2024 share, followed by Commercial & Professional Services, Media & Entertainment, Semiconductors & Semiconductor Equipment, and Technology

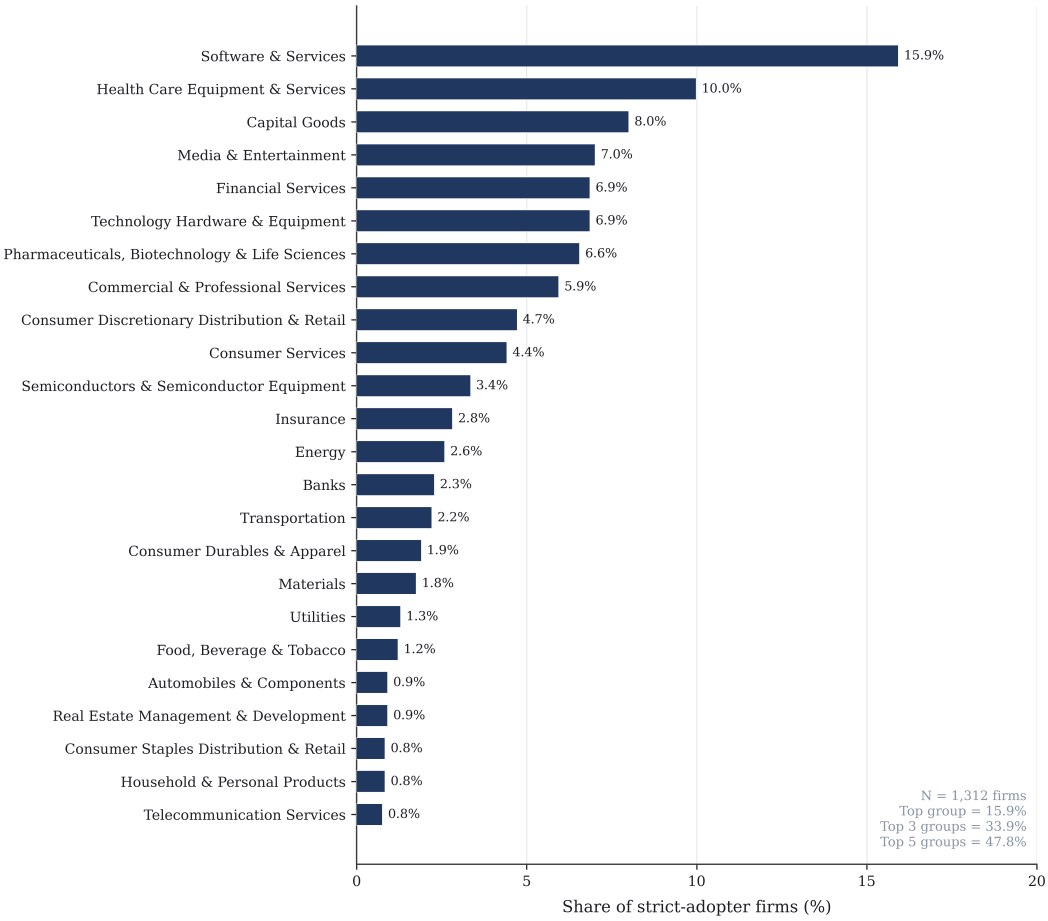


Figure IA.6: **Industry composition of firms with strict AI-deployment disclosures.** The figure reports each four-digit GICS industry group's share of unique firms with at least one deduplicated corporate transcript event classified as strict AI deployment ($AID = 1$). Each firm is counted once and assigned to its modal GICS industry group during 2011–2024, with ties resolved using the most recently observed classification. Percentages are computed relative to all strict-adopter firms with an available industry-group assignment. The dashed vertical line denotes the share implied by an equal distribution across the represented industry groups.

Hardware & Equipment. Panel B reports the 2024 cross-section. Its overall benchmark of 38.8 percent conditions on firms observed in the 2024 panel and therefore differs from the full-universe cumulative share reported in Figure 1.

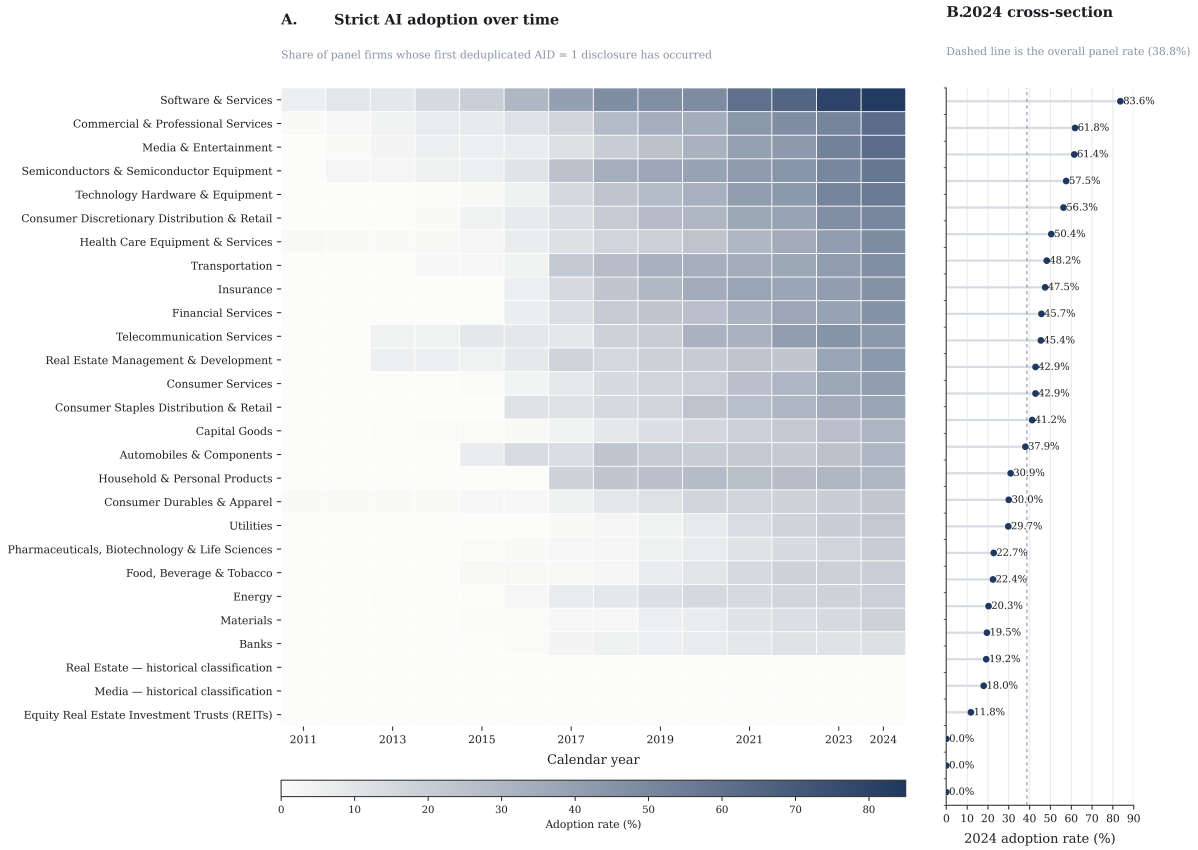


Figure IA.7: **Strict AI adoption by GICS industry group and year.** Panel A reports, for each four-digit GICS industry group and calendar year, the share of observed panel firms whose first deduplicated strict AI-deployment disclosure ($AID = 1$) has occurred by that year. Adoption is treated as an absorbing firm-level stock: once a firm enters the strict adopter group, it remains classified as an adopter in subsequent observed years. Each industry-year denominator is the number of firms from that industry group observed in the panel in that year. Industry groups are ordered by their 2024 adoption rates. Panel B reports the corresponding 2024 cross-section; the dashed vertical line denotes the strict-adoption rate among all firms observed in the 2024 panel.

IA.4.3 Strict generative-AI adoption by industry and year

Figure IA.8 reports the analogous diffusion patterns for strict generative-AI deployment. A firm enters the GenAI adopter stock following its first event classified as strict generative-AI deployment ($GenAID = 1$). In contrast to the broader AI measure, GenAI diffusion is sharply back-loaded: adoption is limited before the final years of the sample and increases

substantially in 2023–2024.

The diffusion wave is also more concentrated across industries. Software & Services exhibits the highest 2024 adoption rate, with substantial uptake in Media & Entertainment and Commercial & Professional Services. Several other information-intensive industries display meaningful but lower penetration, while adoption remains limited in much of the remaining cross-section. Panel B compares these endpoint rates with the overall GenAI adoption rate in the 2024 panel.

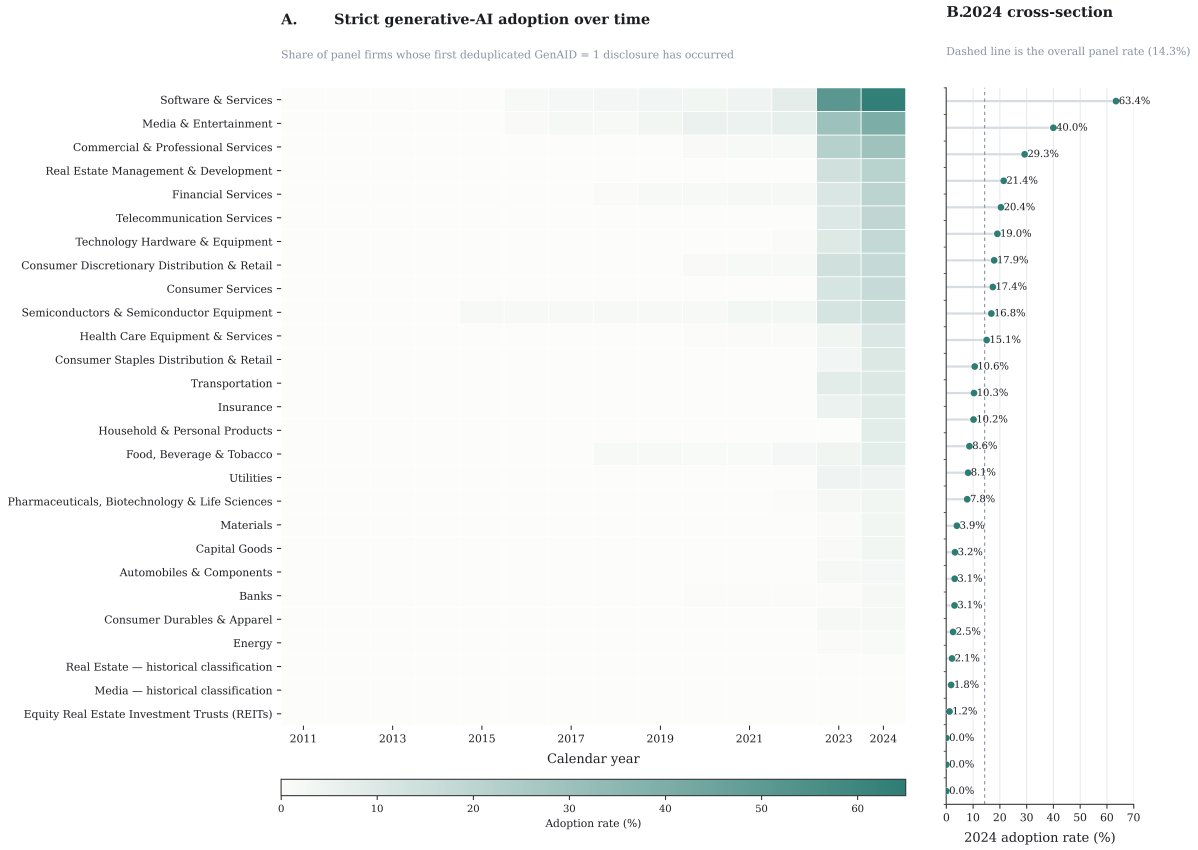


Figure IA.8: **Strict generative-AI adoption by GICS industry group and year.** Panel A reports, for each four-digit GICS industry group and calendar year, the share of observed panel firms whose first deduplicated strict generative-AI deployment disclosure ($GenAID = 1$) has occurred by that year. Adoption is treated as an absorbing firm-level stock, and each industry-year denominator is the number of firms from that industry group observed in the panel in that year. Industry groups are ordered by their 2024 adoption rates. Panel B reports the corresponding 2024 cross-section; the dashed vertical line denotes the strict GenAI adoption rate among all firms observed in the 2024 panel.

IA.4.4 Diffusion paths across levels of industry aggregation

Figure IA.9 examines whether the broad acceleration and cross-industry heterogeneity in strict AI disclosures depend on the level of GICS aggregation. It plots the cumulative share of observed firms that have made a strict deployment disclosure at the six-digit industry, four-digit industry-group, and two-digit sector levels.

Across all three levels, strict disclosures remain uncommon early in the sample and increase substantially after the mid-2010s. Meaningful dispersion remains visible by 2024 both across narrow industries and across broad sectors. The pattern therefore is not solely an aggregate calendar-time phenomenon and supports measuring the competitive state at the industry level.

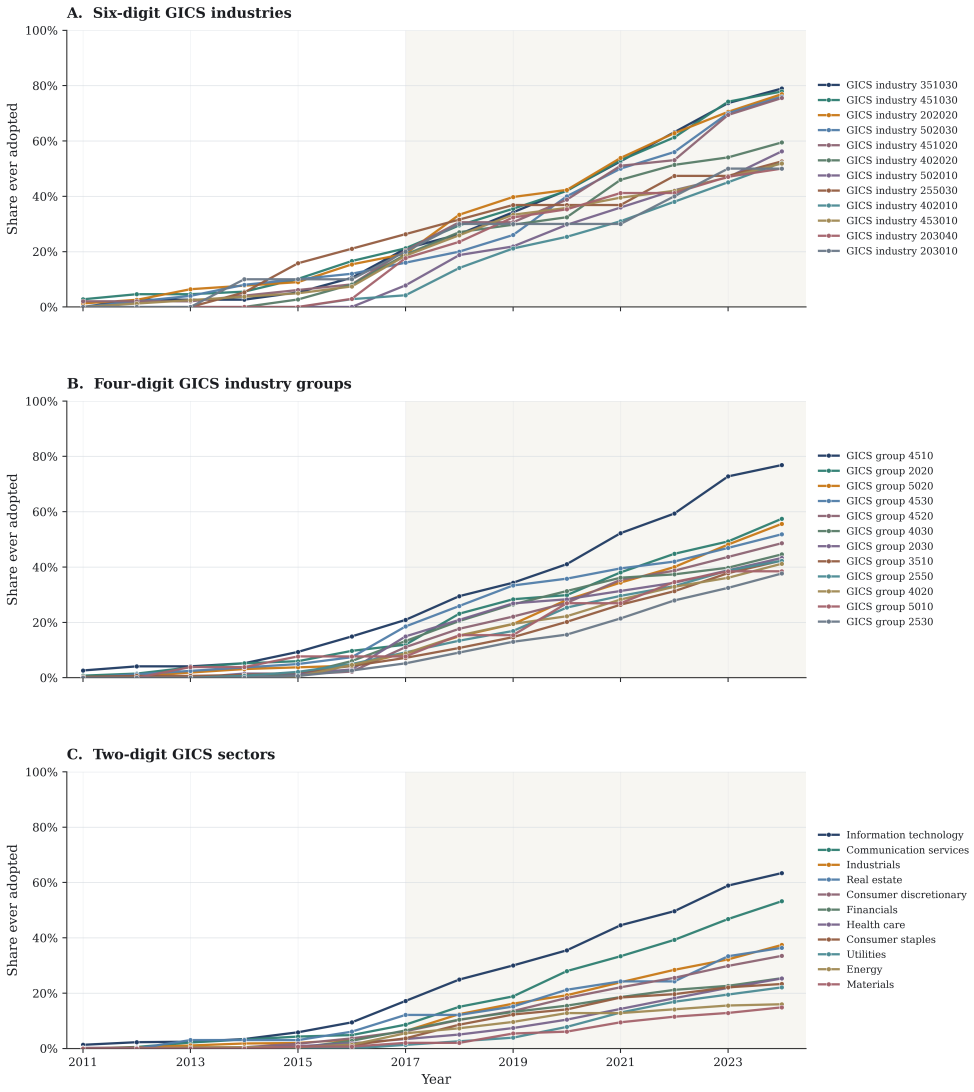


Figure IA.9: **Strict AI deployment-disclosure paths across industry aggregations.** The panels report the annual share of observed firms whose first strict AI-deployment disclosure has occurred by that year, separately for six-digit GICS industries, four-digit GICS industry groups, and two-digit GICS sectors. Firm-level disclosure status is absorbing, while aggregate shares also reflect annual panel composition. The displayed six-digit industries and four-digit groups contain at least ten sample firms; the text should state the exact additional selection rule used to choose the displayed categories. Panel C reports all sufficiently populated sectors. The shaded period beginning in 2017 is a descriptive visual reference and is not the treatment definition used in the main analysis.

IA.5 Model derivations and proofs

This appendix provides the assumptions and derivations underlying the model in Section 3.

IA.5.1 Regularity conditions

Assumption 1 (State processes)**.** The diffusion law $\dot{S} = m(S)$ has a unique solution on $[0, 1]$, with m Lipschitz, $m(S) \geq 0$, and $m(1) = 0$. It is order preserving: $S_0' \geq S_0$ implies $S_t' \geq S_t$ for every t. The application-profitability state follows Equation (2), with $\sigma_X > 0$, under the valuation measure.

Admissible policies satisfy the integrability conditions required for the objectives below and the transversality condition

$$\lim_{T\to\infty} \mathbb{E}\big[e^{-rT}\,\big|V^A(a, K_T, S_T, X_T)\big|\big] = 0,$$

together with the analogous condition for V^N before deployment. The discount rate r is sufficiently large relative to payoff growth and μ_X for the value functions to be finite.

Assumption 2 (Technology and costs)**.** The functions f, g, L, F, and K_0 are twice continuously differentiable and satisfy

$$f' > 0, \qquad f'' < 0, \qquad g > 0, \qquad g' < 0, \qquad L' > 0, \qquad F' < 0, \qquad K_0' \geq 0.$$

The post-deployment control problem is concave, admits an optimal Markov policy, and has a value function that is sufficiently smooth wherever derivatives are used.

Assumption 3 (Monotone deployment)**.** Let the deployment action be $d \in \{0, 1\}$, where $d = 0$ denotes continuation and $d = 1$ denotes immediate deployment. The advantage of immediate deployment over continuation is nondecreasing in the application-profitability state X and has increasing differences in capability a and the deployment action. The transition law of (S, X) is independent of a, and the continuation operator preserves these orderings.

Assumption 4 (Decreasing differences after deployment)**.** The post-deployment Bellman operator preserves decreasing differences in capability a and diffusion S. The primitive flow payoff provides the economic basis for this restriction because

$$\frac{\partial^2 \pi^A(a, K, S, X)}{\partial a \, \partial S} = X g'(S) f(K) < 0.$$

The analysis maintains that this ordering is preserved through the endogenous complementary-investment response.

The monotone-deployment and decreasing-differences assumptions are maintained shape restrictions rather than estimated results. Assumption 3 delivers a deployment policy that is monotone in application profitability and capability. Assumption 4 extends the negative capability–diffusion interaction in current flow profits to the entire post-deployment value function.

IA.5.2 Post-deployment investment

Let $\mathcal{A}(a, K, S, X)$ denote the set of admissible nonnegative investment policies. Policies are progressively measurable, satisfy the integrability and transversality conditions required for the objective to be finite, and preserve $K_t > 0$.

Post-deployment value is defined in Equation (6). For completeness, its control representation is

$$V^A(a, K, S, X) = \sup_{\{I_t\}_{t \geq 0} \in \mathcal{A}(a,K,S,X)} \mathbb{E}_{K,S,X} \left[\int_0^\infty e^{-rt} \left\{ \pi^A(a, K_t, S_t, X_t) - C(I_t, K_t) \right\} dt \right],$$

subject to

$$dK_t = (I_t - \delta K_t)\, dt, \qquad dS_t = m(S_t)\, dt, \qquad dX_t = \mu_X X_t \, dt + \sigma_X X_t \, dW_t, \qquad I_t \geq 0,$$

with initial state

$$(K_{t=0}, S_{t=0}, X_{t=0}) = (K, S, X).$$

Expectations are taken under the valuation measure.

For a sufficiently smooth function $h(K, S, X)$, define the state generator

$$\mathcal{L}h = m(S)h_S + \mu_X X h_X + \frac{1}{2}\sigma_X^2 X^2 h_{XX}. \tag{IA.23}$$

IA.5.2.1 Derivation of the post-deployment HJB

Let $\mathcal{A}_h(a, K, S, X)$ denote the set of admissible investment policies over the interval $[0, h]$. For any sufficiently small $h > 0$, the dynamic-programming principle gives

$$V^A(a, K, S, X) = \sup_{\{I_t\}_{0\le t\le h}\in\mathcal{A}_h} \mathbb{E}_{K,S,X}\Bigg[\int_0^h e^{-rt}\left\{\pi^A(a, K_t, S_t, X_t) - C(I_t, K_t)\right\} dt + e^{-rh}V^A(a, K_h, S_h, X_h)\Bigg].$$

Applying Itô's lemma to $V^A(a, K_t, S_t, X_t)$ gives

$$\mathbb{E}[dV^A] = \left[V_K^A(I_t - \delta K_t) + \mathcal{L}V^A\right] dt.$$

Using $e^{-rh} = 1 - rh + o(h)$, subtracting the current value $V^A(a, K, S, X)$ from both sides, dividing by h, and letting $h \downarrow 0$ yields

$$rV^A = \pi^A + \max_{I\ge 0}\left\{-C(I, K) + V_K^A(I - \delta K)\right\} + \mathcal{L}V^A.$$

Substituting

$$C(I, K) = I + \frac{\gamma}{2}\frac{I^2}{K}$$

gives Equation (7).

Proposition 1 (Irreversible investment rule)**.** Under Assumptions 1 and 2, the optimal gross investment policy satisfies

$$\frac{I^*(a, K, S, X)}{K} = \max\left\{0, \frac{V_K^A(a, K, S, X) - 1}{\gamma}\right\},$$

as stated in Equation (8). The firm undertakes no gross investment when $V_K^A \leq 1$ and invests at a rate increasing in marginal $q^K = V_K^A$ when $V_K^A > 1$.

Proof. The terms in Equation (7) that depend on I are

$$\Phi(I) = -I - \frac{\gamma}{2}\frac{I^2}{K} + V_K^A I.$$

Because

$$\Phi''(I) = -\frac{\gamma}{K} < 0,$$

the objective is strictly concave in I. The unconstrained first-order condition is

$$-1 - \gamma\frac{I}{K} + V_K^A = 0.$$

At an interior solution,

$$\frac{I}{K} = \frac{V_K^A - 1}{\gamma}.$$

Imposing $I \geq 0$ through the Kuhn–Tucker conditions gives

$$\frac{I^*}{K} = \max\left\{0, \frac{V_K^A - 1}{\gamma}\right\}.$$

□

Proposition 2 (Fixed-state target)**.** Freeze S and X and suppose the optimum is interior at a stationary point. Then $K^*(a, S, X)$ satisfies Equation (9); equivalently,

$$aXg(S)f'(K^*(a, S, X)) = (r + \delta)(1 + \gamma\delta) - \frac{\gamma}{2}\delta^2 \equiv \chi > 0.$$

Moreover,

$$K_a^* > 0, \qquad K_X^* > 0, \qquad K_S^* < 0.$$

The proposition is a fixed-state comparative-static device. Along the actual path, both S_t and X_t evolve, so $K^*(a, S_t, X_t)$ is a moving target rather than a literal steady state.

Proof. Freeze S and X and consider the associated deterministic post-deployment control

problem. Define the current-value Hamiltonian

$$\mathcal{H}^C = \pi^A(a, K, S, X) - I - \frac{\gamma}{2}\frac{I^2}{K} + q^K(I - \delta K),$$

where

$$q^K = V_K^A$$

is the current-value costate, or equivalently the shadow value of one additional unit of installed complementary capital.

The current-value costate equation is

$$\dot{q}^K = rq^K - \frac{\partial \mathcal{H}^C}{\partial K}.$$

Holding the control I fixed when taking the Hamiltonian derivative,

$$\frac{\partial \mathcal{H}^C}{\partial K} = aXg(S)f'(K) + \frac{\gamma}{2}\left(\frac{I}{K}\right)^2 - \delta q^K.$$

Consequently,

$$\dot{q}^K = (r + \delta)q^K - aXg(S)f'(K) - \frac{\gamma}{2}\left(\frac{I}{K}\right)^2.$$

At an interior stationary target,

$$\dot{K} = 0 \qquad \text{and} \qquad \dot{q}^K = 0.$$

The capital law implies

$$I = \delta K, \qquad \text{so that} \qquad \frac{I}{K} = \delta.$$

The interior investment first-order condition from Proposition 1 gives

$$q^K = 1 + \gamma\frac{I}{K} = 1 + \gamma\delta.$$

Substituting into the costate equation yields

$$aXg(S)f'(K^*) = (r+\delta)(1+\gamma\delta) - \frac{\gamma}{2}\delta^2 \equiv \chi.$$

Notice that

$$\chi = r + \delta + \gamma r\delta + \frac{\gamma}{2}\delta^2 > 0.$$

Define

$$H(a, S, X, K) = aXg(S)f'(K) - \chi.$$

At the target,

$$H(a, S, X, K^*) = 0.$$

Because

$$H_K = aXg(S)f''(K) < 0,$$

the implicit-function theorem gives

$$K_a^* = -\frac{H_a}{H_K} = -\frac{Xg(S)f'(K^*)}{aXg(S)f''(K^*)} > 0,$$

$$K_X^* = -\frac{H_X}{H_K} = -\frac{ag(S)f'(K^*)}{aXg(S)f''(K^*)} > 0,$$

and

$$K_S^* = -\frac{H_S}{H_K} = -\frac{aXg'(S)f'(K^*)}{aXg(S)f''(K^*)} < 0.$$

The final inequality follows from $g'(S) < 0$, $f'(K) > 0$, and $f''(K) < 0$. □

IA.5.3 Optimal stopping and the free boundary

For any admissible stopping time τ, define

$$\mathcal{V}_\tau(a, S, X) = \mathbb{E}\left[\int_0^\tau e^{-rt}\pi^N(a, S_t)\,dt + e^{-r\tau}J(a, S_\tau, X_\tau)\right].$$

Then

$$V^N(a,S,X) = \sup_{\tau} \mathcal{V}_\tau(a,S,X).$$

The stopping and continuation sets are

$$\mathcal{S} = \left\{(a,S,X) : V^N(a,S,X) = J(a,S,X)\right\},$$

and

$$\mathcal{C} = \left\{(a,S,X) : V^N(a,S,X) > J(a,S,X)\right\}.$$

Proposition 3 (Variational inequality and boundary conditions)**.** Under Assumptions 1–3, V^N satisfies Equation (12) in the viscosity sense and classically wherever it is smooth. If the stopping set is one-sided in X, it can be written as

$$\mathcal{S} = \left\{(a,S,X) : X \geq X^*(a,S)\right\}.$$

At a regular interior boundary, value matching (14) and smooth pasting (15) hold.

Proof. The dynamic-programming principle over a short interval $\Delta > 0$ gives

$$V^N = \max\left\{J,\ \mathbb{E}\left[\int_0^\Delta e^{-rt}\pi^N(a,S_t)\,dt + e^{-r\Delta}V^N(a,S_\Delta,X_\Delta)\right]\right\}.$$

Expanding the continuation value using Itô's lemma and letting $\Delta \downarrow 0$ yields

$$\pi^N + \mathcal{L}V^N - rV^N = 0$$

in the continuation region. Because immediate deployment is always available,

$$V^N \geq J.$$

Combining the continuation equation with this obstacle condition gives

$$\max\left\{\pi^N + \mathcal{L}V^N - rV^N,\ J - V^N\right\} = 0.$$

Assumption 3 requires that the advantage of deployment over continuation increase with the application-profitability state X. The optimal deployment action is therefore monotone in X. Whenever the stopping set is nonempty and admits an interior boundary, it can be represented as

$$\mathcal{S} = \{(a, S, X) : X \geq X^*(a, S)\}.$$

Value matching follows from continuity of the value function and the availability of immediate deployment. Because X_t is a regular diffusion at an interior boundary with $\sigma_X X^* > 0$, the standard smooth-fit condition gives

$$V_X^N(a, S, X^*) = J_X(a, S, X^*).$$

□

Proposition 4 (Capability and deployment timing)**.** Under Assumption 3, the stopping set is increasing in capability: if

$$(a, S, X) \in \mathcal{S}$$

and $a' \geq a$, then

$$(a', S, X) \in \mathcal{S}.$$

Hence, wherever the one-sided representation applies,

$$X^*(a', S) \leq X^*(a, S).$$

Proof. Consider the finite-horizon Bellman recursion with actions "wait" and "deploy." Assumption 3 imposes increasing differences between capability and the deployment action. The transition law of (S, X) is independent of a, so the continuation operator preserves increasing differences. The pointwise maximum over actions preserves this order. Backward induction therefore gives a deployment policy monotone in a for every finite horizon. Discounting and the growth bounds imply convergence to the infinite-horizon

value, preserving monotonicity of the stopping set. A one-sided stopping set then implies the stated boundary ordering. □

IA.5.4 Diffusion, defensive deployment, and conditional value

The current operating gain from being deployed rather than remaining undeployed, holding K fixed, is

$$\Delta\pi(a, K, S, X) = L(S) + aXg(S)f(K). \tag{IA.24}$$

Its diffusion derivative is

$$\Delta\pi_S = L'(S) + aXg'(S)f(K). \tag{IA.25}$$

The second term is negative and captures the erosion of differentiation. The first is positive and captures the competitive loss avoided by deploying.

Assumption 5 (Defensive dominance in the late region)**.** On a region $\mathcal{D}$ relevant for late adopters, the advantage of immediate deployment over continuation is nondecreasing in diffusion S. A transparent local force supporting this condition is

$$L'(S) > -aXg'(S)f(K),$$

but because deployment changes future investment and continuation values, the maintained condition is the present-value ordering rather than the flow inequality alone.

Assumption 5 is a maintained ordering condition on the late-adoption region, not a global implication of the primitive flow-payoff restrictions. The local inequality provides economic intuition, but the proposition below requires the corresponding present-value ordering.

Proposition 5 (Defensive deployment)**.** Under Assumptions 1, 2, 3, and 5, the stopping set expands with diffusion on $\mathcal{D}$. Hence, wherever the stopping set is represented by $X \geq X^*(a, S)$,

$$S' \geq S \quad \Longrightarrow \quad X^*(a, S') \leq X^*(a, S) \qquad \text{on } \mathcal{D}.$$

At the same time, the differentiating flow $aXg(S)f(K)$ strictly decreases with S.

Proof. Assumption 5 imposes a monotonicity condition between diffusion and the deployment action. Applying the same monotone dynamic-programming argument used in Proposition 4, an increase in S cannot switch the optimal action from deploy to wait within $\mathcal{D}$. Thus, the stopping set expands with S, and the one-sided boundary in X weakly falls. The final statement follows directly from $g'(S) < 0$. □

Proposition 6 (Conditional value compression)**.** Under Assumptions 1 and 2, post-deployment value is weakly decreasing in diffusion:

$$V_S^A(a, K, S, X) \leq 0.$$

Under Assumption 4, it also has decreasing differences in capability and diffusion:

$$V_{aS}^A(a, K, S, X) \leq 0.$$

Consequently, the scarcity value in Equation (17) satisfies

$$\mathcal{R}(a, K, S, X) \geq 0, \qquad \mathcal{R}(a, K, 1, X) = 0, \qquad \mathcal{R}_S(a, K, S, X) \leq 0.$$

Proof. Take two initial diffusion states $S' \geq S$. By the order-preserving diffusion law, the path beginning from S' remains weakly above the path beginning from S at every subsequent date.

For any fixed feasible investment policy, $g' < 0$ implies that the discounted operating payoff is weakly lower along the higher-diffusion path. Diffusion enters the flow payoff but does not alter the law of motion for K, the nonnegativity constraint on investment, or the admissibility conditions. The feasible control set is therefore identical across the two initial diffusion states. Taking the supremum over that common policy set preserves the payoff ordering:

$$V^A(a, K, S', X) \leq V^A(a, K, S, X).$$

This establishes $V_S^A \leq 0$ wherever the value function is differentiable.

The primitive deployed payoff has cross-partial

$$\frac{\partial^2 \pi^A}{\partial a \, \partial S} = X g'(S) f(K) < 0.$$

Assumption 4 requires the post-deployment Bellman operator to preserve this decreasing-differences property through the endogenous investment response. Its fixed point therefore satisfies

$$V_{aS}^A \leq 0.$$

Because $S \leq 1$ and V^A is weakly decreasing in S, Equation (17) implies

$$\mathcal{R}(a, K, S, X) \geq 0.$$

At $S = 1$, the two value terms in the definition coincide, so $\mathcal{R}(a, K, 1, X) = 0$. Finally,

$$\mathcal{R}_S(a, K, S, X) = V_S^A(a, K, S, X) \leq 0.$$

□

The defensive-deployment and conditional-value-compression propositions jointly deliver the combination used in the empirical interpretation. Higher diffusion can move laggards toward deployment because remaining behind becomes increasingly costly, while reducing the value of an already deployed firm's differentiating capability.

IA.5.5 Average q is not the scarcity value

Let B_t denote the value of baseline operations and non-AI assets, let $\Omega(K_t)$ denote the replacement value of installed AI-related capital, and let $\mathcal{G}_t$ collect growth and adjustment-option value. A schematic market-value decomposition is

$$M_t = B_t + \Omega(K_t) + \mathcal{G}_t + \mathcal{R}_t.$$

An empirical average-q ratio such as

$$q_t^{\text{avg}} = \frac{M_t}{K_t^P + K_t^I}$$

does not identify $\mathcal{R}_t$ without additional structure. Capitalizing intangibles corrects an important denominator problem, but the residual valuation gap can still contain growth options, adjustment effects, and other economic rents. The empirical capitalization exercise should therefore be described as rejecting the narrow omitted-intangible-capital explanation, not as directly measuring scarcity rents.

IA.5.6 Mapping to the empirical design

Let τ_i^A denote the model's internal deployment stopping time and let τ_i^D denote the first strict public disclosure providing evidence of deployment. The two dates need not coincide. Disclosure may occur after implementation begins and before the technology is fully integrated. The empirical mapping is therefore:

- The underlying empirical event is τ_i^D, the first strict public disclosure. For annual outcomes, that date is subsequently aligned to fiscal periods as described in Section 4.1.
- Capability a is proxied by pre-disclosure constructed intangible capital, R&D intensity, and related readiness measures.
- Diffusion S is proxied by the predetermined, lagged, leave-one-out industry disclosure share.
- Complementary capital K is proxied by constructed stocks of knowledge and organization capital. Post-disclosure changes in those stocks, R&D expenditure, and booked intangibles provide evidence on complementary accumulation.
- Peters–Taylor total q is a capital-adjusted valuation outcome, not a direct observation of $\mathcal{R}$.

The positive association between lagged peer diffusion and first strict disclosure is a reduced-form counterpart of the defensive-deployment mechanism. The framework does not by itself sign the market reaction at the public disclosure date. At an interior internal-deployment boundary, value matching implies that the firm is indifferent between immediate deployment and continued waiting. A disclosure-window valuation response additionally depends on market expectations and on the information conveyed by the disclosure. Accordingly, a flat or negative response among late disclosers is consistent with limited remaining differentiation value and defensive implementation, but it is not a direct structural prediction of the stopping model.

IA.6 Risk-set, mover-definition, and support robustness

The first-mover-minus-follower contrast is not an artifact of transcript observability or imprecise disclosure timing. Figure IA.10 compares the baseline raw-q contrast with three alternative samples defined in Appendix A.1: a current-period-observable sample, a right-censored sample, and a timing-restricted sample that retains only adopters whose disclosure interval is no more than one year.

For raw q, the early $[+1, +2]$ first-mover-minus-follower difference is $+0.447$ in the baseline sample ($p = 0.001$), $+0.307$ in the currently-observable sample ($p = 0.011$), $+0.310$ in the right-censored sample ($p = 0.011$), and $+0.461$ when the disclosure gap is restricted to one year ($p = 0.005$). The corresponding $[+4, +5]$ differences are $+0.192$, $+0.184$, $+0.159$, and $+0.280$, respectively, and none is statistically distinguishable from zero. Thus both the observability restrictions and the tighter disclosure-timing restriction preserve the early-positive, later-imprecise pattern.

IA.6.1 Minimum industry size and complete-window support

The baseline first-mover classification uses the bottom quartile of the predetermined diffusion-state distribution among adopters in the firm's industry. To verify that the result is not driven by quartiles formed in industries with few adopters, I repeat the

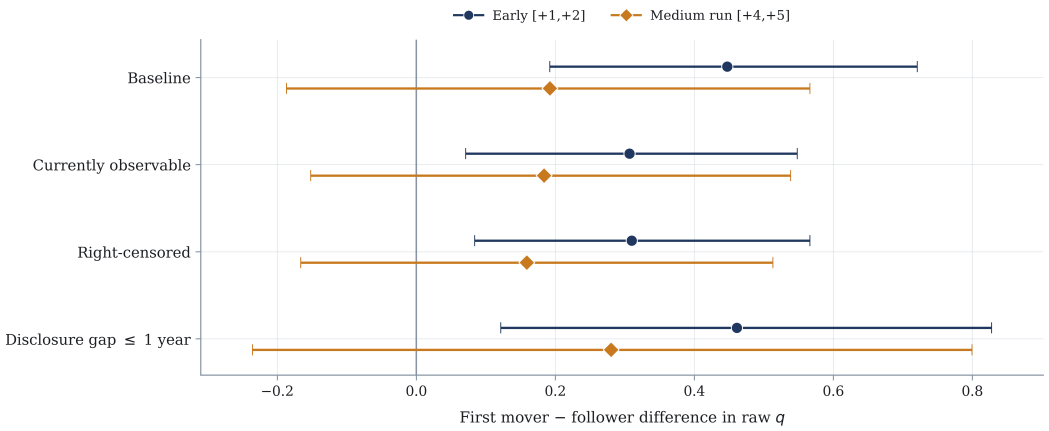


Figure IA.10: **Observation and timing robustness of the first-mover valuation premium.** The figure reports first-mover-minus-follower differences in raw Tobin's q over the early window $[+1, +2]$ (circles) and medium-run window $[+4, +5]$ (diamonds). The first three rows vary transcript observability: the baseline, current-period-observable, and right-censored samples. The final row retains all never-treated firms but restricts adopters to those whose absolute interval between the last observed non-strict disclosure and first strict disclosure is no more than 365 days. Horizontal bars are 95% percentile confidence intervals from 999 firm-cluster bootstrap replications of the untreated-outcome imputation estimator.

classification after requiring at least eight adopters per industry. I also examine whether the any-observation support rule affects the window estimates. Under the baseline rule, a firm contributes to a window when it has at least one nonmissing treated outcome within that window. The complete-window specification instead requires the outcome to be observed at both event years in the indicated two-year window.

Table IA.9 reports the results. Requiring at least eight adopters per industry produces estimates that are nearly identical to the baseline. Over $[+1, +2]$, the first-mover-minus-follower differences are 0.444 for raw book-assets q, 0.378 for intangible-adjusted q, and 1.181 for Peters–Taylor total q. All three are statistically significant. The corresponding $[+4, +5]$ estimates are substantially smaller and statistically indistinguishable from zero.

The complete-window restriction reduces the number of contributing firms, particularly among followers at the later horizon, but preserves the economic conclusion. The early differences remain positive and statistically significant at 0.372, 0.322, and 0.993 across the three valuation perimeters. At $[+4, +5]$, the complete-window estimates are -0.043, 0.002, and -0.214, respectively, and none is statistically distinguishable from zero. The early valuation differential is therefore not generated by small-industry quartile assignments or by firms contributing only a single observation to a two-year window.

Table IA.9: Mover-definition and complete-window robustness

Outcome	Robustness	$[+1,+2]$		$[+4,+5]$	
		Difference (p)	N_{FM}/N_F	Difference (p)	N_{FM}/N_F
Raw book-assets q	Minimum 8 adopters	0.444 (0.004)	317/619	0.154 (0.450)	312/247
	Complete window	0.372 (0.010)	350/503	−0.043 (0.826)	328/166
Intangible-adjusted q	Minimum 8 adopters	0.378 (0.001)	317/619	0.119 (0.297)	312/247
	Complete window	0.322 (0.001)	350/503	0.002 (0.992)	328/166
Peters–Taylor total q	Minimum 8 adopters	1.181 (0.001)	277/557	0.152 (0.559)	271/213
	Complete window	0.993 (0.002)	306/447	−0.214 (0.404)	286/144

Notes: Entries are first-mover-minus-follower differences from the untreated-outcome imputation estimator; centered firm-cluster-bootstrap p-values are in parentheses. N_{FM}/N_F reports the number of contributing first movers and followers. The minimum-industry-size specification applies the within-industry first-mover classification only in industries containing at least eight adopters and otherwise uses the baseline any-observation support rule. The complete-window specification retains the baseline mover classification but requires a nonmissing imputed treatment effect at every event year in the indicated two-year window: both $k = 1$ and $k = 2$ for $[+1,+2]$, or both $k = 4$ and $k = 5$ for $[+4,+5]$. Inference uses 999 firm-cluster bootstrap replications, re-estimating the untreated firm and year effects in each draw. The corresponding baseline any-observation estimates are reported in Table 2.

IA.7 Depth of AI engagement and the valuation response.

Section 5.5 argues that the revaluation tracks concrete deployment rather than AI discussion. This appendix makes that a dose-response. I sort firms into three tiers by the maximum AI-adoption score they ever reach in their transcripts, confirmed deployment (maximum AID = 1), pilot-stage exploration (0maximum AID = 0.5), and mention-only (maximum AID = 0 among firms that match the AI-keyword filter), and estimate the aggregate Sun–Abraham valuation path for each tier around the firm's first transcript in that tier. All three tiers are compared with the identical control pool of transcript-universe firms that never match the AI-keyword filter, so the only object that changes across panels is the intensity of the firm's AI activity, not the benchmark. Figure IA.11 plots the three paths. The medium-run revaluation of raw Tobin's q is large and precisely estimated for confirmed deployment, weaker and less precise for pilot-stage exploration, and flat for mention-only firms. The valuation response rises monotonically across the stages of AI engagement, which is what a return to deployment predicts and what mere AI discussion does not.

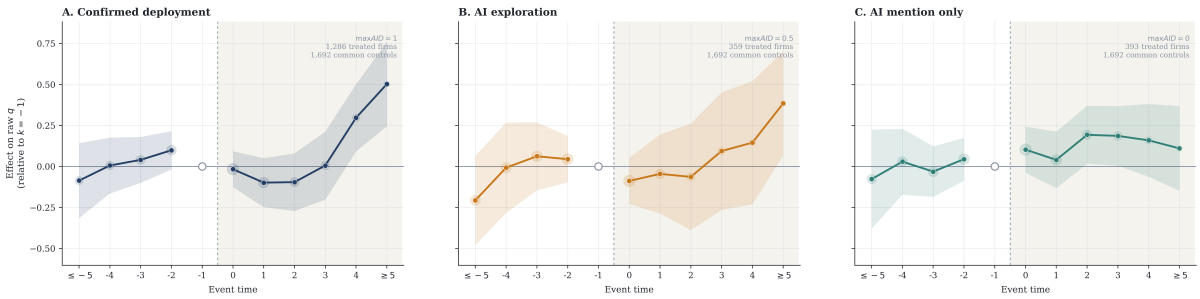


Figure IA.11: **The valuation response rises with the depth of AI engagement.** Aggregate interaction-weighted Sun–Abraham paths for raw Tobin's q around a firm's first transcript in each intensity tier: confirmed deployment (maximum AID = 1; panel A), pilot-stage exploration (maximum AID = 0.5; panel B), and mention-only (maximum AID = 0 among keyword-matched firms; panel C). All three panels use the identical control pool of transcript-universe firms that never match the AI-keyword filter (1,692 firms), a cleaner "no AI activity" benchmark than the never-strict-adopter pool used in the main event studies. The revaluation is large and precise for confirmed deployment, weaker for pilot-stage exploration, and flat for mention-only firms; the monotonic ordering, not the level of any single tier, is the object of interest. The pilot and mention tiers are imprecisely estimated. Event time is relative to $k = -1$; endpoints pool $k \leq -5$ and $k \geq 5$; shaded regions are 95% confidence intervals; standard errors clustered by firm.

IA.8 Short-window market response to disclosed AI deployment

This section uses short-window abnormal returns to establish the market relevance of the first strict deployment disclosure and to examine whether the immediate response varies with predetermined diffusion. These tests complement the annual valuation evidence but are not used as the paper's primary diffusion design.

IA.8.1 Abnormal returns around the first strict disclosure

I match each adopter's first strict disclosure date to daily CRSP returns and estimate abnormal returns using the Fama–French five-factor model. Table IA.10 reports group-specific announcement returns. Both groups experience positive average abnormal returns around the event. For first movers, mean CAR$[-1, +1]$ is 0.64% and mean CAR$[0, +1]$ is 0.56%; for followers, the corresponding estimates are 0.58% and 0.44%. Thus the dated events are not arbitrary transcript timestamps: they coincide, on average, with economically meaningful market reactions.

The adjusted first-mover-minus-follower coefficient for CAR$[0, +1]$ is positive, at 137

basis points, but its industry-clustered standard error is 115 basis points ($p = 0.239$). Figure IA.12 also shows that the cumulative paths begin to separate before day zero. That pattern can reflect anticipation, leakage, other information released before the call, or imperfect alignment between the transcript timestamp and the first moment at which investors learned the information. I therefore use these returns to establish market relevance, not to identify the early-mover differential.

Table IA.10: Abnormal returns around the first strict AI-deployment disclosure

Group	Window	N	Mean CAR (%)	Patell Z	BMP t
Panel A. Group-specific abnormal returns					
First movers	$[-1,+1]$	358	0.6419	2.9626	1.6746
First movers	$[0,0]$	358	−0.2146	−1.2748	−0.9846
First movers	$[0,+1]$	358	0.5614	3.1107	1.5080
Followers	$[-1,+1]$	765	0.5766	3.3734	1.5589
Followers	$[0,0]$	765	0.4968	4.3629	2.0254
Followers	$[0,+1]$	765	0.4402	3.6705	1.2304
Panel B. Regression-adjusted first-mover-minus-follower difference					
Outcome		Effect (bps)	SE (bps)	t	p
CAR$[0,+1]$		137.03	115.14	1.19	0.239

Notes. Panel A reports mean five-factor-adjusted cumulative abnormal returns around each firm's first strict deployment disclosure. Patell and BMP statistics standardize event returns using the pre-event estimation window. Panel B reports the coefficient on the first-mover indicator from the disclosure-level regression with event-month and industry effects and pre-event firm controls; standard errors are clustered by industry. The adjusted coefficient need not equal the raw difference between the two group means.

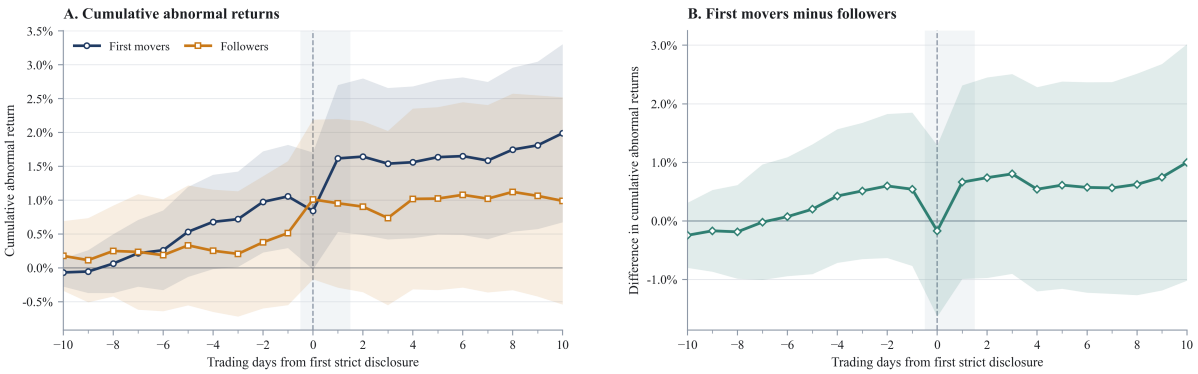


Figure IA.12: **Cumulative abnormal returns around the first strict deployment disclosure.** Panel A plots group-specific cumulative abnormal returns from ten trading days before through ten trading days after the first strict disclosure. Panel B plots the first-mover-minus-follower difference. The vertical dashed line marks the event date. Both groups have positive short-window reactions, but the paths begin to separate before day zero; the figure is therefore descriptive and is not used as identification of a disclosure-day first-mover effect.

IA.8.2 Diffusion-conditioned announcement returns and statistical power

I next regress CAR$[0, +1]$ on either the continuous diffusion state or the first-mover indicator. The continuous coefficient is negative and the first-mover coefficient is positive, the signs predicted by the scarcity mechanism, but neither is precisely estimated. With event-month and industry effects and pre-event controls, a one-standard-deviation increase in diffusion is associated with a 45-basis-point smaller announcement return ($p = 0.280$), while the first-mover coefficient is 108 basis points ($p = 0.278$). More saturated fixed effects yield similar qualitative conclusions.

The imprecision is economically informative. At conventional size and 80% power, the minimum detectable effect is approximately 117 basis points for the continuous-diffusion design and 275 basis points for the first-mover indicator. These thresholds are substantially larger than the raw difference between the group-specific mean returns in Table IA.10. The short-window test therefore cannot distinguish zero from a moderate diffusion effect; it should not be read as evidence that announcement responses are invariant to diffusion.

Table IA.11: Diffusion-conditioned announcement returns and minimum detectable effects

Design	Specification	Effect (bps)	SE (bps)	p	MDE (bps)	N
Continuous diffusion	Month + industry FE, controls	−45.45	41.67	0.280	116.67	990
First mover indicator	Month + industry FE, controls	107.55	98.17	0.278	274.89	990
Continuous diffusion	Fine industry + quarter + content FE	−27.55	101.31	0.787	283.68	990
First mover indicator	Fine industry + quarter + content FE	124.81	124.97	0.322	349.91	990

Notes. The continuous-diffusion coefficient is the effect of a one-standard-deviation increase in the peer-deployment state; a negative sign means smaller announcement returns at higher diffusion. The first-mover coefficient compares the first-mover group with followers. MDE is the two-sided minimum detectable effect at the 5% significance level and 80% power, evaluated using the reported standard error. Standard errors are clustered by industry.

IA.9 Accounting-perimeter and calibration robustness

Figure IA.13 shows that the first-mover life cycle is robust to how intangible capital is capitalized. Each row is a valuation perimeter or a Peters–Taylor calibration; within each row, the two markers are the first-mover-minus-follower difference at the early window $[+1, +2]$ and at the later window $[+4, +5]$, with horizontal bars the 95% percentile intervals

from 999 firm-cluster bootstrap replications. Reading down the rows, the early difference is positive and significant and the later difference is small on every perimeter (raw, intangible-adjusted, Peters–Taylor) and under all six calibrations: capitalizing intangibles rescales the premium but leaves its early concentration and subsequent attenuation intact.

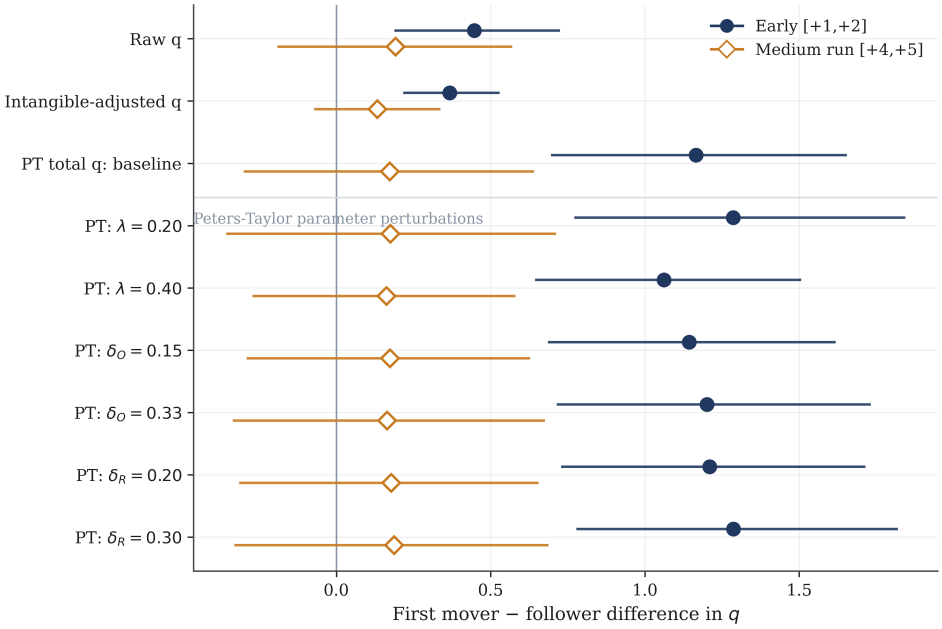


Figure IA.13: **Accounting-perimeter and intangible-capital calibration robustness.** First-mover-minus-follower valuation difference at the early $[+1, +2]$ (left markers) and later $[+4, +5]$ (right markers) windows, for raw q, intangible-adjusted q, Peters–Taylor total q, and six alternative Peters–Taylor calibrations (one per row). Horizontal bars are 95% percentile confidence intervals from 999 firm-cluster bootstrap replications. The early premium is present and the later difference attenuated on every perimeter and calibration.

IA.9.1 Fitted paths at low and high predetermined diffusion

The main text reports the low-minus-high diffusion contrast from the continuous heterogeneous event-study design. Figure IA.14 decomposes that contrast into the two underlying fitted paths. I evaluate equation (29) at the 25th and 75th percentiles of the predetermined diffusion state among adopters, corresponding to peer-deployment shares of approximately 0.147 and 0.468. These values are fixed at the firm's first strict deployment disclosure and do not vary over the subsequent event-time path.

Panel A reports raw book-assets q. The fitted low-diffusion path is generally above

the high-diffusion path from event year zero through year four, but the high-diffusion estimates have wide confidence intervals, the paths converge at the pooled endpoint, and the joint post-disclosure interaction test does not reject ($p = 0.400$). Raw book-assets q therefore provides directional rather than precise evidence of a continuous diffusion gradient.

Panels B and C show a clearer decomposition on the capitalized perimeters. For intangible-adjusted q, the low-diffusion path remains near zero while the high-diffusion path becomes progressively negative from event year zero through year three and remains below the low-diffusion path through year four. The joint post-disclosure interaction test rejects at $p = 0.037$. The Peters–Taylor total-q results are stronger: the low-diffusion path again remains close to zero, while the high-diffusion path falls sharply over event years zero through four. The corresponding joint test rejects at $p = 0.006$.

The fitted paths show that the significant capitalized-q contrast is generated primarily by an unfavorable high-diffusion response rather than by a uniformly large positive low-diffusion effect. This pattern parallels the group-specific evidence in Figure 3, where followers initially decline relative to their counterfactual before recovering at later horizons. The figure is used to interpret the source of the post-disclosure contrast. Formal inference on the difference between the low- and high-diffusion paths uses the simultaneous sup-t bands reported in Figure 2, panel C. Pre-disclosure diagnostics are reported separately in Table C.1.

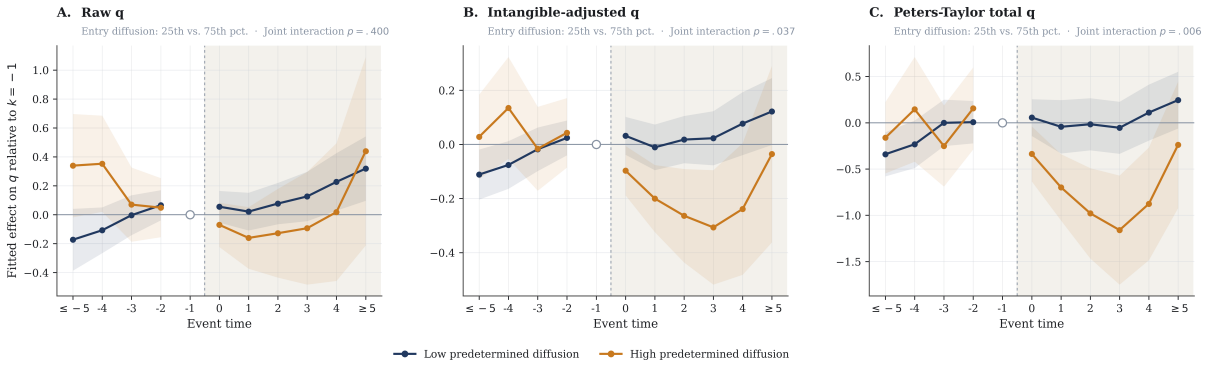


Figure IA.14: **Fitted valuation paths at low and high predetermined diffusion.** The figure evaluates the continuous heterogeneous event-study specification at the 25th and 75th percentiles of the leave-one-out industry diffusion state measured immediately before the firm's first strict deployment disclosure. Panels A, B, and C report raw book-assets q, intangible-adjusted q, and Peters–Taylor total q, respectively. Event time $k = -1$ is omitted, and the endpoints pool $k \leq -5$ and $k \geq 5$. Shaded regions are pointwise 95% confidence intervals based on firm-clustered standard errors. The panel subtitles report joint Wald tests of the null that all six post-disclosure diffusion-interaction coefficients, $k = 0, \ldots, 5$, equal zero. Vertical scales differ across panels because the valuation ratios use different capital perimeters. Formal inference on the low-minus-high difference uses the simultaneous sup-t bands reported in Figure 2, panel C.

IA.9.2 Financial firms, utilities, and the operating-capital perimeter

A physical-plus-intangible replacement-capital denominator is less natural for financial institutions and regulated utilities than for operating firms. I therefore re-estimate the continuous predetermined- diffusion specification after excluding financial firms and utilities. Financial firms are identified by SIC codes 6000–6999 and utilities by SIC codes 4900–4999. Classification is fixed at the firm level using the firm's modal status over the panel.

The nonfinancial, nonutility sample contains 31,927 firm-years from 3,161 firms, including 1,123 strict disclosers with exact stock timing. After requiring a nonmissing valuation outcome and the variables needed for estimation, the raw and intangible-adjusted specifications use 29,428 observations from 3,094 firms. For Peters–Taylor total q, I additionally require gross property, plant, and equipment of at least \$1 million and total gross physical-plus-intangible replacement capital of at least \$1 million. The resulting operating-firm sample contains 27,440 estimation observations from 2,938 firms, including 1,095 treated firms with a predetermined diffusion state.

Table IA.12 reports joint tests of the six post-disclosure diffusion interactions. Removing financial firms and utilities has little effect on the raw book-assets result: the joint p-value changes from 0.286 in the broad sample to 0.272 in the nonfinancial, nonutility sample. Raw q therefore continues to point in the same direction without providing precise joint evidence.

The capitalized measures are different. On the nonfinancial, nonutility sample, the six post-disclosure diffusion interactions are jointly significant for intangible-adjusted q ($p = 0.024$). On the operating-screened Peters–Taylor sample, the joint null is rejected at $p = 0.003$. For both capitalized measures, the interaction coefficient is negative in every event year from $k = 0$ through $k = 4$ and individually significant at conventional levels. The gradient weakens at $k = 5$, where event-time support is lower.

The exclusion exercise is not used to claim uniformly cleaner pre-disclosure paths: the $k = -4$ interaction is positive in the raw and intangible-adjusted specifications. Its implication is instead that the post-disclosure capitalized-q gradient is not an artifact of pooling financial institutions and utilities with operating firms. If anything, the diffusion-dependent valuation pattern is sharper on the economic perimeters for which replacement capital has its most natural interpretation.

Table IA.12: Financial-firm and utility exclusion in the continuous diffusion design

Sample	Outcome	Joint χ^2	p-value	Observations	Firms	Treated firms
Broad sample	Raw book-assets q	7.40	0.286	37,552	3,828	1,293
Nonfinancial, nonutility	Raw book-assets q	7.56	0.272	29,428	3,094	1,123
Nonfinancial, nonutility	Intangible-adjusted q	14.55	0.024	29,428	3,094	1,123
Operating-firm screen	Peters–Taylor total q	19.69	0.003	27,440	2,938	1,095

Notes: The table reports joint Wald tests of $H_0 : \delta_0 = \delta_1 = \cdots = \delta_5 = 0$, where δ_k is the coefficient on event time k interacted with the firm's predetermined leave-one-out industry diffusion state. The diffusion state is measured immediately before the firm's first strict deployment disclosure and is held fixed across its event-time path. Specifications include firm and industry-by-calendar-year fixed effects; standard errors are clustered by firm. Event time is aligned to the first fiscal year-end on or after disclosure, and $k = -1$ is omitted.
Financial firms have SIC codes 6000–6999 and utilities have SIC codes 4900–4999; firm status is fixed using the modal classification over the panel. The operating-firm Peters–Taylor sample additionally requires gross property, plant, and equipment of at least \$1 million and gross physical plus total intangible replacement capital of at least \$1 million. "Treated firms" denotes firms with an exact stock cohort and a nonmissing predetermined diffusion state in the estimation sample.

IA.10 Dynamic peer catch-up within first movers

Section 5.4 reports the pooled association between peer catch-up and valuation over event years one through five. Figure IA.15 reports the underlying event-year coefficients. The estimates are generally negative but individually noisy, which is why the pooled coefficient is the preferred summary in the main text.

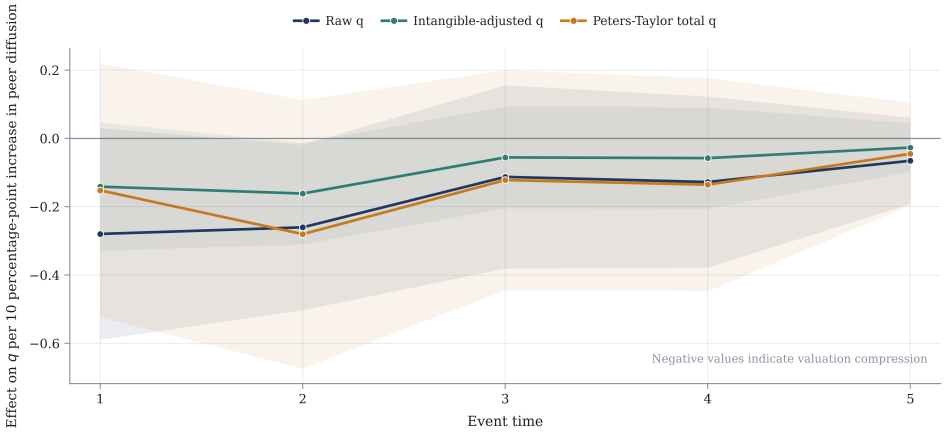


Figure IA.15: **Dynamic association between peer catch-up and valuation.** The figure reports event-year-specific coefficients on subsequent peer catch-up within first movers. Peer catch-up is the current leave-one-out industry diffusion share minus the firm's predetermined diffusion state at its own disclosure. Specifications include firm, calendar-year, and event-time fixed effects. Effects are expressed per ten-percentage-point increase in peer diffusion. Confidence intervals are simultaneous 95% bands.

IA.11 Capability and diffusion: full cross-section

Section 6.2 reports the capability-by-diffusion interaction on Peters–Taylor total q under sector fixed effects. This appendix reports the full cross-section: all three valuation perimeters, both fixed-effect structures, and the post-adoption integration-intensity specification. The specification is the adopter-level regression of the long-run revaluation Δq_i on the predetermined diffusion state $S_i = S^{-i}_{g(i),G_i-1}$, a standardized capability proxy A_i, and their interaction, with adoption-cohort fixed effects and log-asset controls; the Peters–Taylor calibration grid for this interaction is in Appendix D.

Table IA.13 collects the interaction for all three valuation perimeters under both the sector and the fine-industry fixed-effect structures.

Table IA.13: Capability × diffusion interaction, all perimeters and both fixed-effect structures

		Sector FE		Fine-industry FE	
Perimeter	Capability	Coef.	t	Coef.	t
Raw q	K_{intan}/AT	−0.282	−2.43	−0.156	−1.10
	R&D/AT	−0.206	−0.95	−0.031	−0.14
Intangible-adj. q	K_{intan}/AT	−0.166	−2.82	−0.095	−1.31
	R&D/AT	−0.155	−1.70	−0.060	−0.67
Peters–Taylor q	K_{intan}/AT	−0.386	−3.20	−0.159	−1.19
	R&D/AT	−0.583	−4.96	−0.200	−1.37

Notes: The table reports the coefficient on $S_i \times A_i$; continuous regressors are standardized and standard errors are clustered by industry. The interaction is negative across all reported Peters–Taylor parameterizations under sector fixed effects and attenuates materially under fine-industry fixed effects. The latter specification absorbs substantial variation in the fine-industry diffusion state and is reported as a sensitivity.

Integration intensity. Replacing pre-adoption capability with post-adoption AI-integration intensity, the intensity-by-diffusion interaction on Peters–Taylor total q is -0.325 ($t = -3.02$) using mean integration intensity and -0.763 ($t = -5.50$) using cumulative intensity: deeper AI use is valuable primarily when the technology remains scarce. Because integration intensity is measured after adoption, I treat this as descriptive corroboration rather than as a predetermined test.

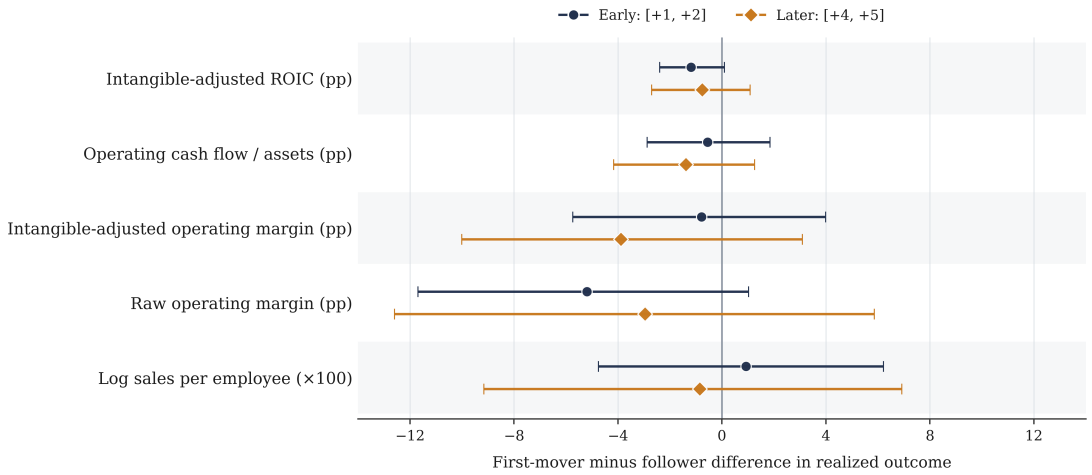
IA.12 Realized operating outcomes: full family and support sensitivity

Section 6.3 reports the two headline operating outcomes. This appendix reports the complete pre-specified family of six outcomes and examines whether the results are sensitive to differences in outcome availability. The accounting constructions and exact definitions of all six outcomes are provided in Section IA.2.3. All estimates are first-mover-minus-follower differences from the untreated-outcome imputation estimator. Inference uses a 999-replication firm-cluster bootstrap that re-estimates the untreated firm and year effects in every draw.

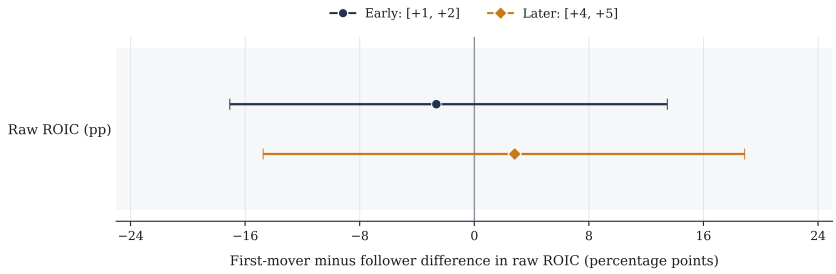
IA.12.1 Full operating-outcome family

Figure IA.16 summarizes the complete outcome family at the two supported post-disclosure windows. Panel A presents five outcomes on a common confidence-interval scale. Panel B presents raw ROIC separately because its confidence intervals are substantially wider.

Nine of the twelve point estimates are negative. None of the positive estimates is statistically precise, and no outcome provides evidence of first-mover operating outperformance within the observed horizon. The largest early negative estimates occur for raw operating margin and intangible-adjusted ROIC, although only the latter approaches conventional significance in the baseline specification.



(a) Five operating outcomes presented on a common confidence-interval scale.



(b) Raw ROIC, presented separately because its confidence intervals are substantially wider.

Figure IA.16: **Realized operating outcomes: first mover minus follower.** The figure reports first-mover-minus-follower differences in imputed treatment effects over event years $[+1, +2]$ and $[+4, +5]$. The six outcomes are defined in Section IA.2.3. Horizontal bars are 95% confidence intervals. The estimator fits untreated firm and year effects using pre-disclosure observations for eventual disclosers and all available observations for never-disclosing firms. Inference uses a 999-replication firm-cluster bootstrap that re-estimates the untreated fixed effects in every draw. First-mover status is the predetermined diffusion classification of Section 4.4.1, while event time is aligned to the first fiscal period lying entirely after disclosure.

Table IA.14 reports the corresponding point estimates and t-statistics. Intangible-adjusted ROIC is 1.18 percentage points lower for first movers in the early window ($t = -1.84$; bootstrap $p = 0.065$). Raw operating margin is also lower early, by 5.19 percentage points, but is less precisely estimated ($t = -1.61$). The remaining estimates are economically and statistically mixed, with no systematic evidence of superior realized performance among first movers.

Table IA.14: Realized operating outcomes: full unbalanced family

	$[+1, +2]$		$[+4, +5]$	
Outcome	Difference	t	Difference	t
Intangible-adjusted ROIC	−1.18	−1.84	−0.76	−0.79
Operating cash flow / assets	−0.55	−0.45	−1.39	−1.01
Raw ROIC	−2.66	−0.32	+2.82	+0.32
Raw operating margin	−5.19	−1.61	−2.96	−0.64
Intangible-adjusted operating margin	−0.78	−0.32	−3.89	−1.16
Labor productivity	+0.93	+0.32	−0.86	−0.21

Notes: The table reports first-mover-minus-follower differences in imputed treatment effects. The six outcomes and their accounting constructions are defined in Section IA.2.3. ROIC, operating-cash-flow, and operating-margin entries are percentage points. Labor productivity is 100 times the log difference in real sales per employee. The unbalanced specification requires an outcome to be observed in at least one year of the indicated post-disclosure window.
Inference uses 999 firm-cluster bootstrap replications that re-estimate the untreated firm and year effects in every draw; all draws were successful for all outcomes. The early intangible-adjusted-ROIC estimate has a centered-bootstrap p-value of 0.065. First-mover status is the predetermined diffusion classification of Section 4.4.1. Event time for these annual flow outcomes begins with the first fiscal period lying entirely after disclosure. Outcome-specific support is approximately 368–380 first movers and 549–591 followers in $[+1, +2]$, and 329–354 first movers and 173–179 followers in $[+4, +5]$.
The operating-margin outcomes require positive sales and assets, sales/assets of at least 0.05, and fiscal-year winsorization at the 2.5th and 97.5th percentiles after applying the denominator screen. Raw ROIC and the operating-margin outcomes have relatively wide confidence intervals but are retained to report the complete pre-specified outcome family.

IA.12.2 Balanced-support and IPCW sensitivity

The unbalanced specification permits a firm to contribute when its outcome is observed in at least one year of the relevant window. Differential outcome availability could therefore affect the first-mover–follower comparison. I assess this possibility in two ways. The balanced-support specification requires the outcome to be observed at every event horizon from $k = -1$ through the end of the relevant post-disclosure window. The IPCW

specification instead reweights the observed outcomes by the estimated inverse probability of remaining observed.

Table IA.15 reports these specifications for the two outcomes emphasized in the main text. The substantive conclusion is unchanged: neither outcome provides evidence of first-mover operating outperformance.

Intangible-adjusted ROIC is negative in both windows under all three specifications. In $[+1, +2]$, the estimated difference is -1.185 percentage points in the unbalanced sample ($p = 0.065$), -0.921 in the balanced sample ($p = 0.171$), and -1.493 under IPCW ($p = 0.046$). Operating cash flow over assets is also negative in every window and specification, although none of those estimates is statistically precise. The later balanced samples are considerably smaller, particularly for followers, but produce the same absence of first-mover outperformance.

Table IA.15: Realized operating outcomes: balanced-support and IPCW sensitivity

Outcome	Specification	$[+1, +2]$		$[+4, +5]$	
		Difference (p)	$N_{\text{FM}}/N_{\text{F}}$	Difference (p)	$N_{\text{FM}}/N_{\text{F}}$
Intangible-adjusted ROIC	Unbalanced	-1.185 (0.065)	370/579	-0.756 (0.417)	344/179
	Balanced	-0.921 (0.171)	367/443	-0.789 (0.495)	312/83
	IPCW	-1.493 (0.046)	370/579	-0.924 (0.341)	344/179
Operating cash flow / assets	Unbalanced	-0.554 (0.655)	380/591	-1.390 (0.300)	354/179
	Balanced	-0.204 (0.866)	377/448	-1.311 (0.486)	322/83
	IPCW	-0.580 (0.660)	380/591	-1.553 (0.281)	354/179

Notes: Entries are first-mover-minus-follower differences in percentage points. The underlying outcome definitions are reported in Section IA.2.3. Centered-bootstrap p-values are reported in parentheses. $N_{\text{FM}}/N_{\text{F}}$ denotes the numbers of contributing first movers and followers.
The unbalanced specification requires at least one observed treated outcome in the indicated window. The balanced specification requires the outcome to be observed at every event horizon from $k = -1$ through the end of the indicated window. IPCW uses stabilized inverse-probability-of-observation weights clipped to $[0.10, 10]$.
All specifications use the untreated-outcome imputation estimator and 999 firm-cluster bootstrap draws that re-estimate the untreated firm and year effects in every replication; all draws were successful. First-mover status uses the predetermined stock-aligned disclosure classification. Event time for these annual flow outcomes begins with the first fiscal period lying entirely after disclosure.

Taken together, the full outcome family and the support sensitivities show that the early valuation differential is not accompanied by superior realized operating performance among first movers within the available horizon. The realization evidence therefore does not identify the source of the valuation differential; its interpretation comes from the capitalization, diffusion-gradient, capability, and peer-catch-up evidence developed

elsewhere in the paper.

IA.13 Labor-share dynamics

This section examines how measured firm-level labor share changes after the first strict AI-deployment disclosure and whether the response differs between first movers and followers. The analysis uses the three firm-level labor-share ratios defined in Section IA.2.4. Measure A divides a micro-based firm-level wage bill by adjusted value added. Measure B uses reported labor expense when available and 0.70 times SG&A otherwise, divided by the same adjusted value-added denominator. Measure C divides the Measure-B labor-expense numerator by operating income before depreciation plus labor expense. Measures A and B share a denominator, while Measures B and C share a numerator. Agreement across the three measures therefore establishes robustness to alternative labor-expense and value-added constructions.

First-mover status is held fixed using the paper's predetermined, stock-aligned diffusion classification. Because labor share is an annual flow outcome, event time is re-aligned to the first fiscal period lying entirely after the strict disclosure. This convention yields a maximum flow-timed sample of 382 first movers and 764 followers. Each mover group is compared with the same pool of 2,581 never-disclosing firms, although usable support varies across the three outcomes and event-time horizons.

I use two complementary estimators. Figure IA.17 reports separate interaction-weighted Sun–Abraham event-study paths for first movers and followers relative to the common never-disclosing comparison pool. Table IA.16 reports direct first-mover-minus-follower differences over the pre-specified $[+1, +2]$ and $[+4, +5]$ windows using the untreated-outcome imputation estimator. Inference for the window contrasts uses a 999-replication firm-cluster bootstrap that re-estimates the untreated firm and year effects in every draw.

Across the three constructions, the first-mover paths remain close to zero during the first several post-disclosure years and decline modestly at the longer supported horizons. The follower paths instead become progressively more negative after disclosure. The

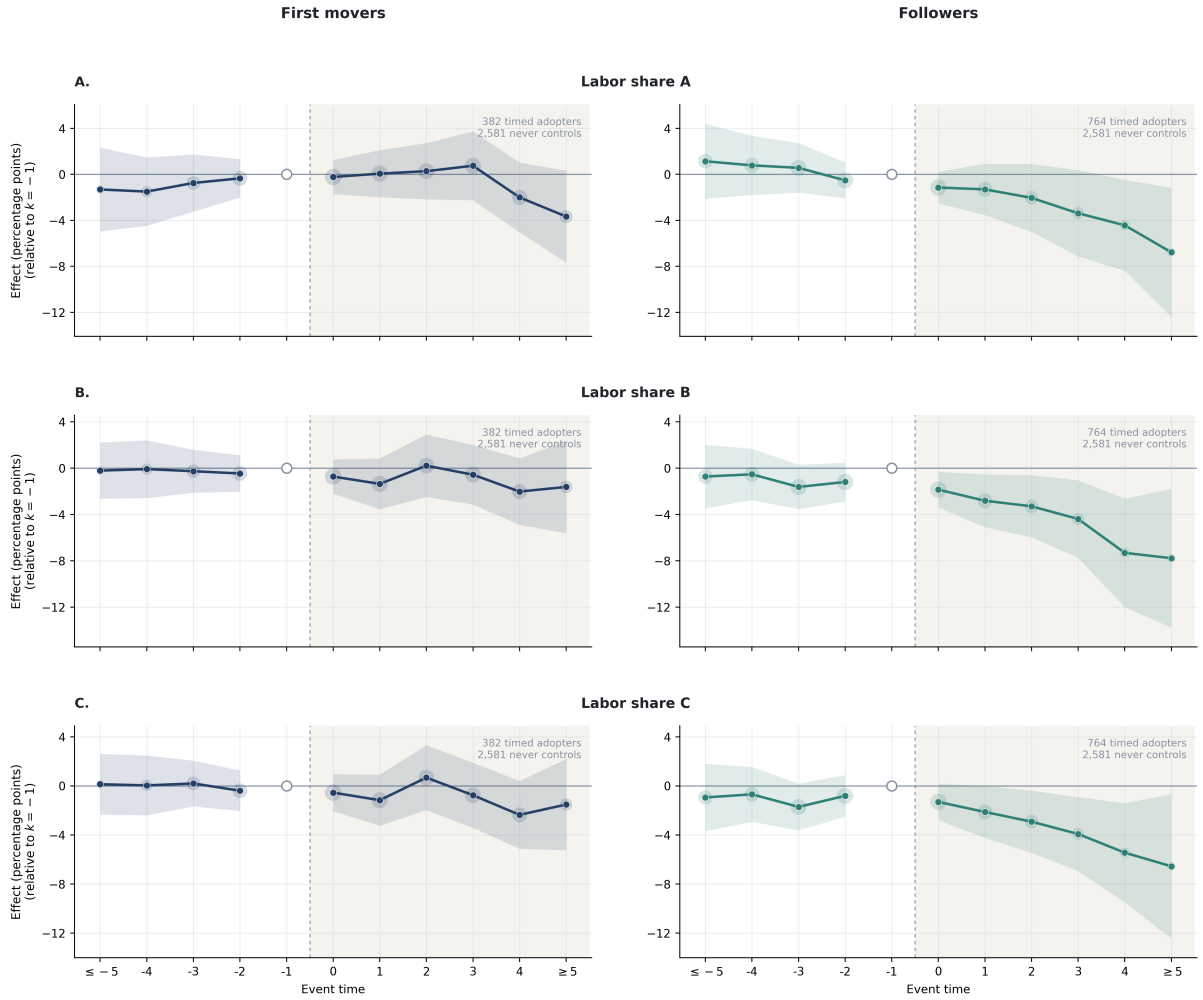


Figure IA.17: **Labor-share dynamics after disclosed AI deployment.** The figure reports interaction-weighted Sun–Abraham event-study paths for three firm-level labor-share measures, separately for first movers (left) and followers (right) against the common never-disclosing comparison pool. Panels A–C correspond to Labor Shares A–C defined in Section IA.2.4. First-mover status uses the predetermined diffusion classification, while event time is aligned to the first fiscal period lying entirely after disclosure. Event time $k = -1$ is omitted, endpoints pool $k \leq -5$ and $k \geq 5$, and shaded regions are 95% confidence intervals based on firm-clustered standard errors.

common ordering across the three measures shows that the larger later decline among followers is not specific to either the micro-based wage bill or a particular value-added denominator.

The direct window comparisons are reported in Table IA.16. In the early $[+1, +2]$ window, the estimated first-mover-minus-follower differences range from 1.67 to 1.76 percentage points and are not statistically precise. In the later $[+4, +5]$ window, all three differences are positive because the measured labor-share decline is larger among followers. The difference is 3.47 percentage points for Measure A ($p = 0.211$), 6.26 points for Measure B ($p = 0.019$), and 4.58 points for Measure C ($p = 0.045$).

Table IA.16: Labor-share responses: first movers versus followers

Outcome	Window	First movers	Followers	FM − follower	p-value	$N_{\mathrm{FM}}/N_{\mathrm{F}}$
Labor share A	$[+1, +2]$	$+0.65$ (1.27)	-1.06 (1.26)	$+1.71$ (1.77)	0.344	316/448
	$[+4, +5]$	-0.61 (1.80)	-4.08 (2.30)	$+3.47$ (2.74)	0.211	302/146
Labor share B	$[+1, +2]$	-0.68 (1.17)	-2.44 (1.10)	$+1.76$ (1.58)	0.256	324/482
	$[+4, +5]$	-1.48 (1.48)	-7.74 (2.17)	$+6.26$ (2.50)	0.019	310/153
Labor share C	$[+1, +2]$	-0.53 (1.14)	-2.20 (1.04)	$+1.67$ (1.50)	0.257	308/480
	$[+4, +5]$	-1.55 (1.41)	-6.13 (1.96)	$+4.58$ (2.30)	0.045	294/157

Notes: Entries are percentage points. The first-mover and follower columns report average imputed treatment effects relative to each group's untreated counterfactual; the difference column reports first mover minus follower. Bootstrap standard errors are in parentheses, and the reported p-values test the direct group difference. Labor Shares A–C are defined in Section IA.2.4. Event time is aligned to the first fiscal period lying entirely after disclosure. Inference uses 999 firm-cluster bootstrap replications that re-estimate the untreated firm and year effects in every draw. The final column reports the numbers of contributing first movers and followers.

The later-window evidence shows that measured labor-share declines are concentrated among followers. The first-mover-minus-follower difference is statistically significant for Measures B and C and positive but imprecisely estimated for Measure A. The common ordering across all three constructions indicates a more pronounced labor-share adjustment among firms disclosing when peer deployment is more widespread.

These firm-level estimates do not separately identify whether the adjustment operates through employment, wages, value added, operating surplus, or changes in produc-

tion composition. The underlying workforce adjustment is analyzed using worker- and occupation-level evidence in our companion labor-market paper.

IA.14 Complementary investment and financing diagnostics

This section collects the secondary investment and financing evidence that supports Section 7. All analyses use the predetermined first-mover classification and the timing convention appropriate to each outcome: fiscal-year-end timing for balance-sheet stocks and the first fully post-disclosure fiscal period for annual flows. I first report the constructed intangible-capital stock, then net equity issuance, and finally the stock-timed debt, cash, and secured-debt contrasts.

IA.14.1 Constructed intangible-capital stock

Figure IA.18 reports the constructed knowledge-and-organization-capital stock over assets. The measure accumulates capitalized R&D and the capitalizable share of SG&A using the perpetual-inventory procedure described in Appendix IA.2.1. The post-disclosure path rises, but the asset-scaled series also exhibits substantial pre-disclosure movement and unstable endpoint behavior. Moreover, the positive first-mover-minus-follower difference in the unscaled dollar stock partly reflects the larger pre-disclosure size of first movers.

IA.14.2 Net equity issuance after disclosed AI deployment

I next examine whether the lower debt response of first movers is accompanied by external equity financing. The preferred measure is constructed from the cash-flow statement as

$$\text{NetEquityIssuance}_{it} = \frac{\texttt{sstk}_{it} - \texttt{prstkc}_{it}}{\texttt{at}_{i,t-1}},$$

where `sstk` is proceeds from stock issuance, `prstkc` is purchases of common and preferred stock, and beginning assets are required to come from the immediately preceding fiscal

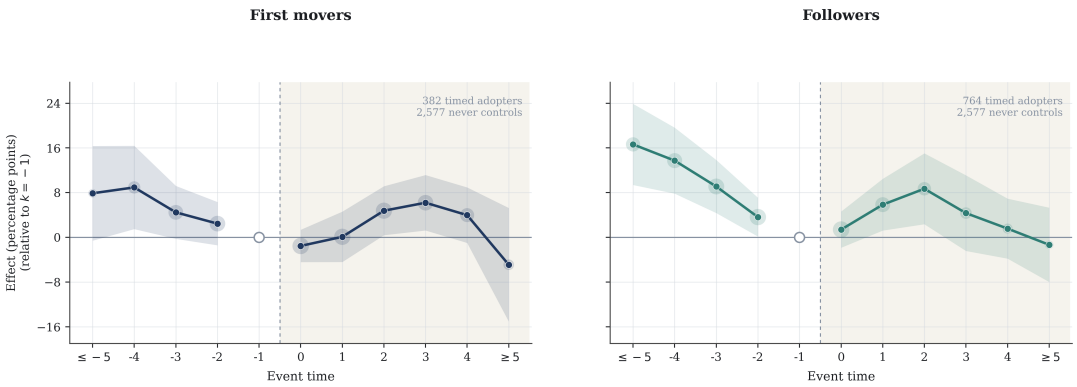


Figure IA.18: **Constructed intangible-capital stock after disclosed AI deployment.** The figure reports interaction-weighted Sun–Abraham paths for the constructed intangible-capital stock over assets, separately for first movers and followers. The stock combines capitalized R&D and organization capital constructed from SG&A. Event time uses the first fully post-disclosure fiscal period. Shaded regions are 95% confidence intervals. Because the asset-scaled series exhibits pre-disclosure movement and the unscaled stock is strongly related to firm size, the figure is interpreted as descriptive rather than as a formal first-mover investment-intensity result.

year. Both cash-flow components must be observed. The measure is winsorized within fiscal year at the first and ninety-ninth percentiles. Negative values denote net repurchases. First-mover status is fixed using the paper's predetermined diffusion classification, whereas event time uses the first fiscal period lying entirely after disclosure because net issuance is an annual financing flow.

Figure IA.19 reports the dynamic post-disclosure paths. Neither group exhibits an increase in net external equity issuance. First movers are modest net repurchasers through the first several post-disclosure years, whereas followers display substantially larger net repurchases, especially in event years zero through two.

Table IA.17 reports the corresponding direct first-mover-minus-follower contrasts from the untreated-outcome imputation estimator. Over $[+1, +2]$, net equity issuance is -6.73 percentage points for first movers and -21.05 percentage points for followers, producing a first-mover-minus-follower difference of $+14.32$ percentage points (95% CI $[7.78, 21.84]$, $p = 0.001$). Over $[+4, +5]$, the group-specific estimates are -1.05 and -3.13 percentage points, respectively, and the difference narrows to $+2.08$ percentage points (95% CI $[-4.06, 8.60]$, $p = 0.504$). The equity-financing difference is therefore concentrated in the early post-disclosure window.

The result narrows the interpretation of the balance-sheet evidence. First movers do

not finance the intangible installation through net external equity issuance. Followers instead combine higher debt with larger net repurchases in the early window, indicating that their debt increase is part of a broader capital-structure and payout adjustment rather than simply external financing of AI investment.

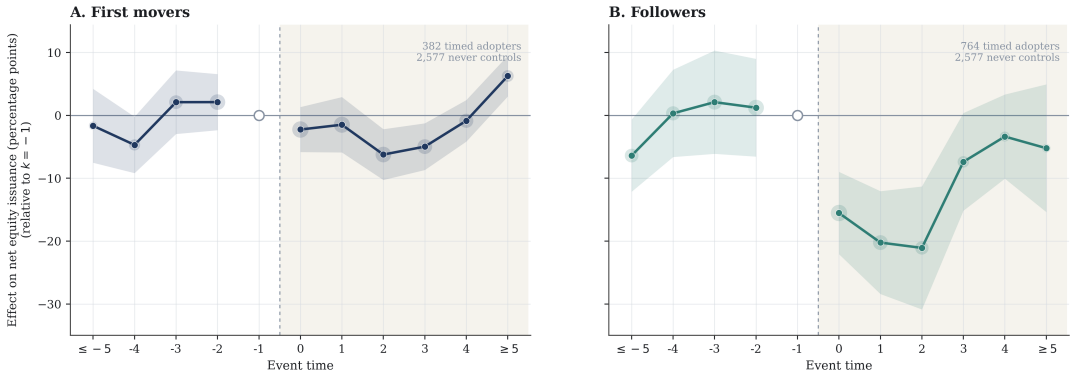


Figure IA.19: **Net equity issuance after disclosed AI deployment.** The figure reports interaction-weighted Sun–Abraham event-study estimates of net equity issuance over beginning-of-period assets, measured as (`sstk` − `prstkc`)/`lagged assets`, separately for first movers and followers against the common never-discloser pool. First-mover status uses the predetermined-diffusion classification, while event time uses the first fully post-disclosure fiscal period because net issuance is an annual flow. Negative estimates indicate net repurchases. Event time $k = -1$ is omitted, endpoints pool $k \leq -5$ and $k \geq 5$, and shaded regions are 95% confidence intervals based on firm-clustered standard errors. The measure is winsorized within fiscal year at the first and ninety-ninth percentiles. Some pre-disclosure movement remains, so the paths are interpreted descriptively.

Table IA.17: Net equity issuance: first movers versus followers

Window	First movers	Followers	Difference	95% confidence interval	p-value
$[+1, +2]$	-6.73	-21.05	$+14.32$	$[7.78,\ 21.84]$	0.001
$[+4, +5]$	-1.05	-3.13	$+2.08$	$[-4.06,\ 8.60]$	0.504

Notes: Entries are percentage points. Net equity issuance is (`sstk` − `prstkc`)/`beginning assets`; negative values indicate net repurchases. Differences are first mover minus follower in the imputed treatment effect. Event time uses the first fiscal period lying entirely after the strict deployment disclosure because issuance is an annual flow. Inference uses a 999-replication firm-cluster bootstrap that re-estimates untreated firm and year effects in every draw (seed 20260727); all draws were successful. Support is 343 first movers and 557 followers at $[+1, +2]$, and 317 and 174 at $[+4, +5]$.

IA.14.3 Balance-sheet adjustment: debt, liquidity, and secured debt

Table IA.18 reports direct first-mover-minus-follower contrasts for debt over assets, cash over assets, secured debt over assets, and the secured share of total debt. Because all four outcomes are fiscal-year-end balance-sheet stocks, event time is aligned to the first fiscal year-end on or after the strict deployment disclosure. The estimator fits firm and year effects on untreated observations, imputes the untreated outcome for each treated firm-year, averages the resulting treatment effects within the pre-specified windows, and uses a 999-replication firm-cluster bootstrap for inference.

The strongest difference appears in medium-run book leverage. Over $[+4, +5]$, first movers increase debt over assets by 2.39 percentage points, compared with 6.45 points among followers, a difference of -4.05 points ($p = 0.038$). The early debt difference is smaller and imprecisely estimated, and the cash-over-assets differences are statistically insignificant in both windows. Secured debt over assets displays a medium-run first-mover-minus-follower difference of -2.24 percentage points ($p = 0.066$), whereas the secured share of total debt does not differ significantly across groups. The evidence therefore identifies a smaller medium-run increase in book leverage among first movers, with a similar but less precisely estimated difference in secured debt over assets.

Table IA.18: Debt, liquidity, and secured-debt responses: first movers versus followers

Outcome	Window	First movers	Followers	Difference	95% confidence interval	p-value
Debt / assets	$[+1, +2]$	$+0.82$	$+2.29$	-1.46	$[-3.96,\ 0.85]$	0.236
Debt / assets	$[+4, +5]$	$+2.39$	$+6.45$	-4.05	$[-7.88,\ -0.31]$	0.038
Cash / assets	$[+1, +2]$	-1.18	-1.24	$+0.06$	$[-1.62,\ 1.83]$	0.945
Cash / assets	$[+4, +5]$	-2.17	-0.93	-1.24	$[-3.34,\ 0.73]$	0.235
Secured debt / assets	$[+1, +2]$	-1.51	-0.53	-0.98	$[-2.60,\ 0.61]$	0.232
Secured debt / assets	$[+4, +5]$	-0.97	$+1.27$	-2.24	$[-4.79,\ 0.05]$	0.066
Secured share of total debt	$[+1, +2]$	-4.36	-2.24	-2.12	$[-6.81,\ 2.39]$	0.382
Secured share of total debt	$[+4, +5]$	-1.13	$+1.15$	-2.28	$[-8.40,\ 3.90]$	0.483

Notes: Entries are percentage points. Differences are first mover minus follower in the imputed treatment effect. Because these are fiscal-year-end balance-sheet stocks, event time uses the first fiscal year-end on or after the strict deployment disclosure. The windows are pre-specified, and a firm contributes when it has at least one observed treated outcome within the window. Inference uses a 999-replication firm-cluster bootstrap that re-estimates untreated firm and year effects in every draw; all 999 draws were successful. First-mover status is the predetermined-diffusion classification used throughout the paper.

Figure IA.20 reports the corresponding event-time paths for secured debt over assets

and the secured share of total debt. First movers show no sustained increase in either measure. The follower paths are imprecisely estimated, while the direct medium-run contrast indicates a larger increase in secured debt over assets among followers. Compustat `dm` measures the stock of secured debt outstanding. These estimates therefore describe changes in the composition of outstanding debt, rather than whether the marginal borrowing contract is secured or unsecured.

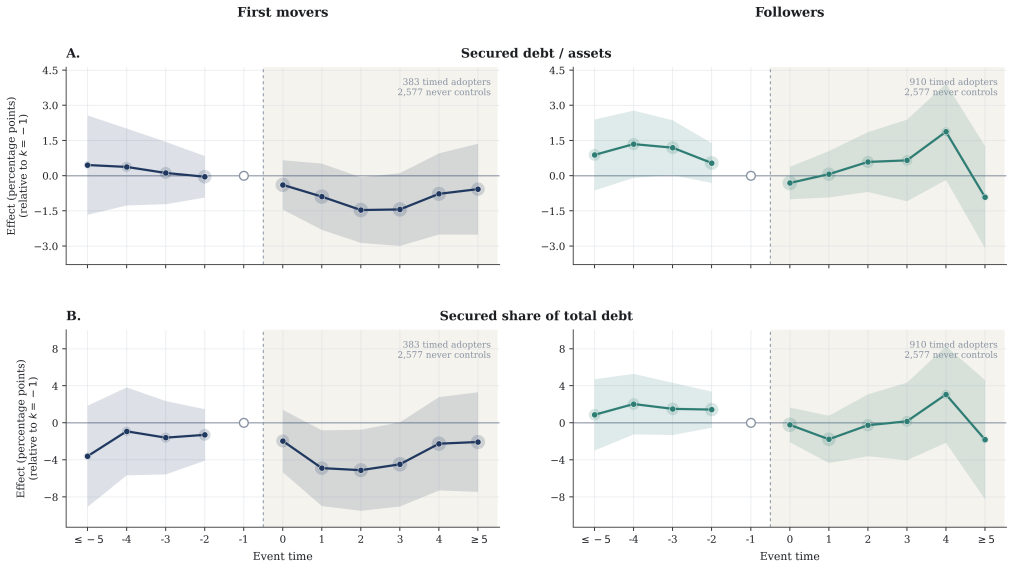


Figure IA.20: **Secured-debt adjustment after disclosed AI deployment.** The figure reports interaction-weighted Sun–Abraham paths for secured debt over assets and the secured share of total debt, separately for first movers and followers against the common never-discloser pool. Secured debt is Compustat `dm`; the secured share is `dm`/(`dltt` + `dlc`). Event time uses the first fiscal year-end on or after disclosure. Event time $k = -1$ is omitted, endpoints pool $k \leq -5$ and $k \geq 5$, and shaded regions are 95% confidence intervals based on firm-clustered standard errors.

IA.15 Cost shares, gross margins, and markups

This appendix examines the operating cost-structure changes that accompany the complementary intangible investment documented in Section 7.1. The estimates use the interaction-weighted Sun–Abraham framework and the predetermined first-mover–follower classification. The gross-margin measure and the three production-based markup constructions are defined in Section IA.2.5. Figures IA.21 and IA.22 report the event-time paths.

IA.15.1 Operating margins and cost shares

Figure IA.21 reports the conventional operating margin, cost of goods sold over sales, and SG&A over sales. Among first movers, both cost shares rise at several post-disclosure horizons, while the operating margin declines. The joint pattern is consistent with software, data, and organizational investment being expensed during installation rather than immediately appearing as capital on the balance sheet. Some pre-disclosure movement and wide confidence intervals remain, so these estimates characterize the accompanying cost structure rather than the model's structural adjustment-cost function.

IA.15.2 Production-based markups and gross margins

Figure IA.22 reports the baseline production-based markup, the markup excluding financial firms, the finance-adjusted markup, and gross profit over sales. Across all four constructions, first movers exhibit no sustained post-disclosure increase in measured price–cost margins. The consistency across accounting and production-based measures indicates that the valuation response is not accompanied by an immediate expansion in measured markups or gross margins.

The production-based measures combine an estimated output elasticity with the flexible-input revenue share, while gross margin provides a direct accounting measure. Their agreement makes the result robust to the choice of margin construction.

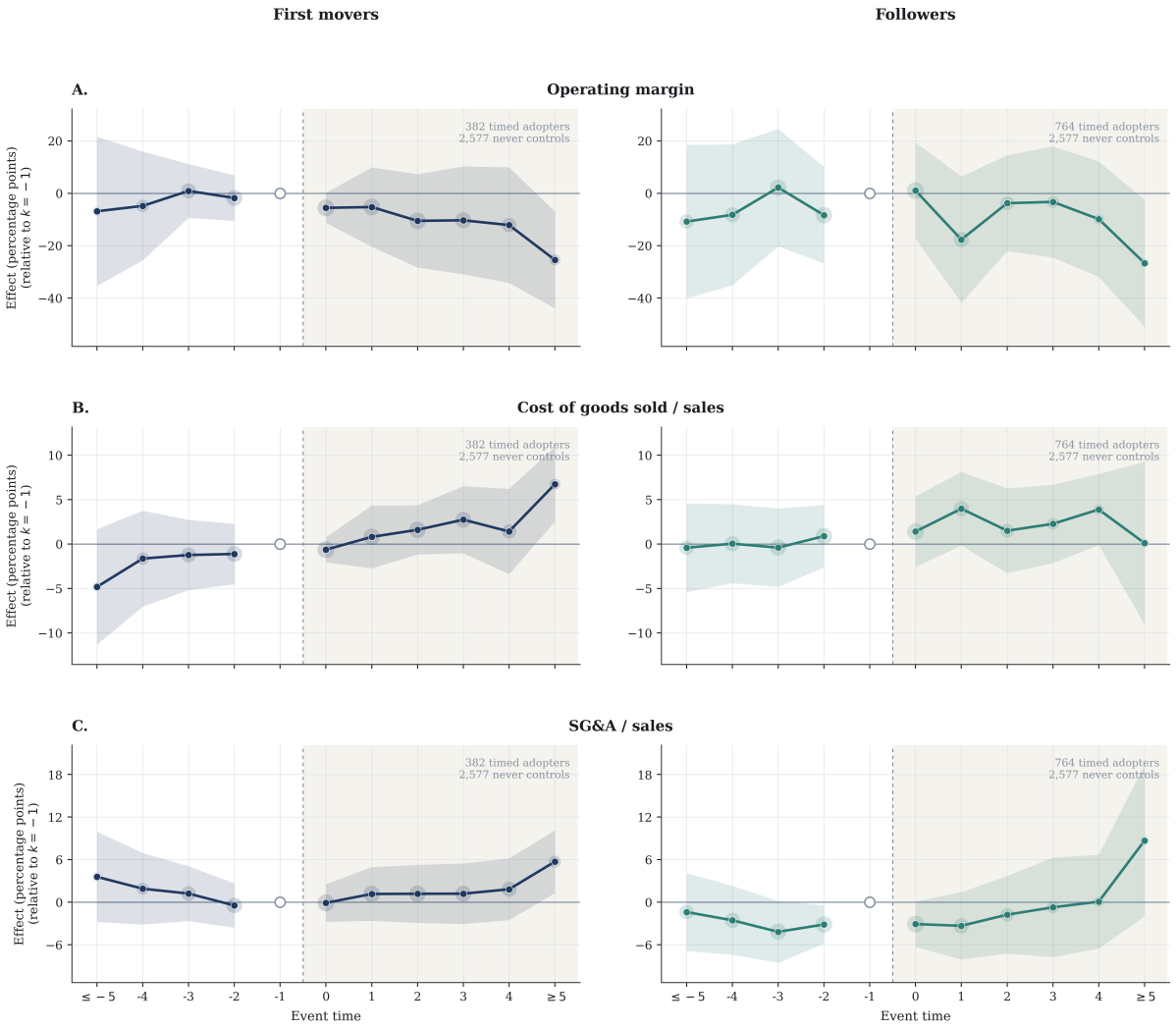


Figure IA.21: **Operating margins and cost shares after disclosed AI deployment.** The figure reports interaction-weighted Sun–Abraham event-study paths for the operating margin (Panel A), cost of goods sold over sales (Panel B), and SG&A over sales (Panel C). Within each panel, the left plot reports first movers and the right plot reports followers, each against the common never-adopter comparison pool. Event time is aligned to the first fiscal period lying entirely after disclosure, and first-mover status uses the predetermined diffusion classification. Event time $k = -1$ is omitted, endpoints pool $k \leq -5$ and $k \geq 5$, and shaded regions are 95% confidence intervals based on firm-clustered standard errors. The intangible-adjusted operating margin is omitted because small denominators generate extreme ratio values.

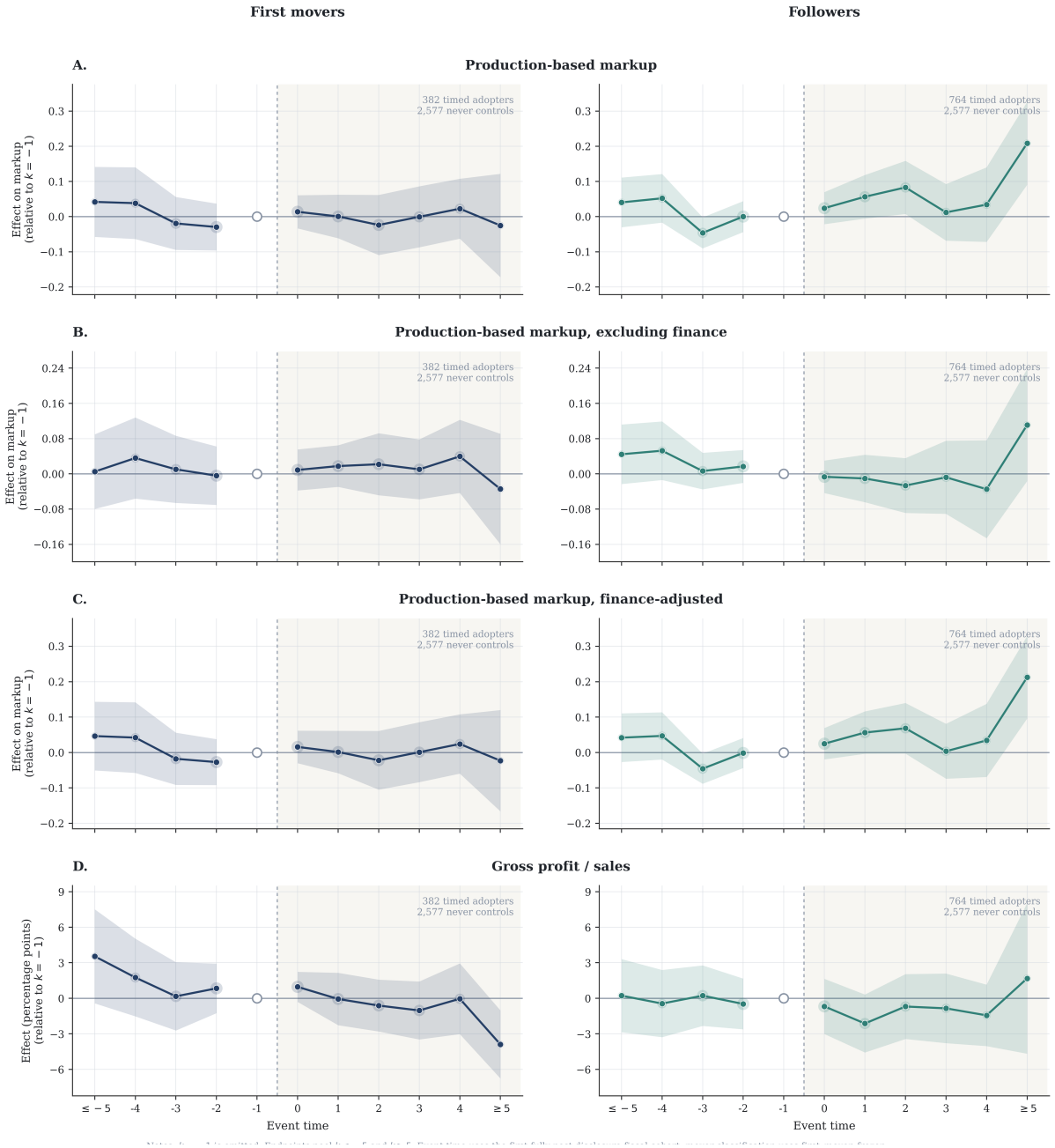


Figure IA.22: **Markup and gross-margin dynamics after disclosed AI deployment.** Panels A–C report interaction-weighted Sun–Abraham event-study paths for the baseline production-based markup, the markup excluding financial firms, and the finance-adjusted markup. Panel D reports gross profit over sales. The four measures are defined in Section IA.2.5. Within each panel, the left plot reports first movers and the right plot reports followers, each against the common never-adopter comparison pool. Event time is aligned to the first fiscal period lying entirely after disclosure, and first-mover status uses the predetermined diffusion classification. Event time $k = -1$ is omitted, endpoints pool $k \leq -5$ and $k \geq 5$, shaded regions are 95% confidence intervals, marker area reflects treated-firm support, and standard errors are clustered by firm.