When Firms Adopt AI Matters: Diffusion, Intangible Investment, and Firm Value — Evidence from Corporate Disclosures

Parand Akbari[1]

Tepper School of Business, Carnegie Mellon University

Abstract

Where do the gains from AI accrue inside firms, and how does the market price them as adoption diffuses? Using 2011–2024 earnings and investor-call transcripts for U.S. public firms, I construct a new firm-level measure of AI deployment (coded as concrete implementation rather than generic discussion) and merge it to Compustat outcomes on costs, factor shares, investment, and valuation. Large language models date each firm's first transition from AI talk to deployment and classify deployments by application type, strategic role, and sourcing mode. I estimate dynamic effects around adoption using staggered-timing event studies and benchmark level effects using diffusion-based panel variation. Three findings emerge. First, adoption primarily reorganizes costs rather than raising pricing power: production-based markup measures are flat while operating cost ratios shift around adoption. Second, adoption is intangible-intensive, coinciding with rising capital/intangibles shares and a declining within-firm labor share. Third, the valuation payoff is sharply decreasing in diffusion: early adopters in low-diffusion industries experience much larger increases in Tobin's q than firms adopting after AI becomes widespread. Overall, the evidence portrays AI as an intangible-heavy technology whose private returns depend critically on adoption timing and competitive diffusion.

Keywords: Artificial intelligence, diffusion, intangibles, Tobin's q, factor shares, corporate disclosures.

JEL: G12, G14, O33, E22, L25.

[1] pakbari@andrew.cmu.edu

1 Introduction

Artificial intelligence (AI) is widely viewed as a general-purpose technology, but its firm-level incidence and economic channels remain unsettled. When firms adopt AI at scale, do the gains materialize as higher markups and market power, as reorganization of operating costs, as a reallocation of value added away from labor toward capital and intangibles, or primarily as a repricing of growth options by equity markets? A second question is about timing: are the gains concentrated among early adopters near the technological frontier, or do they accrue similarly once AI becomes industry "table stakes"?

This paper provides firm-level evidence using a new transcript-based measure of AI deployment (AID). We assemble the near-universe of S&P Capital IQ earnings and investor call transcripts for U.S. public firms from 2011–2024 (187,909 transcript events covering 4,313 firms) and convert these disclosures into a structured panel that dates each firm's first publicly evidenced transition from generic AI discussion to concrete implementation and use. Because simple keyword counts are too noisy for timing and taxonomy—often conflating macro "AI talk" with firm-level deployment and mixing customer adoption with the firm's own adoption—we use a large language model (LLM) with a fixed extraction schema conditioned on the Capital IQ header firm. The LLM returns evidence-gated adoption scores on $\{0, 0.5, 1\}$ with mandatory supporting quotes for auditability and organizes deployment into an interpretable taxonomy (pilot vs. integrated, AI type including Generative AI (GenAI), sourcing mode, and ecosystem role as user vs. vendor). Relative to keyword-frequency measures, this approach improves (i) timing accuracy by dating first deployment at the transcript-event level, (ii) target-company isolation by conditioning extraction on the Capital IQ header firm, (iii) auditability via supporting text, and (iv) economic interpretability via taxonomy (user vs. vendor, in-house vs. external, GenAI vs. non-GenAI).

The paper then links adoption to accounting outcomes, production-based markup measures, factor shares, intangible capital, and market valuation by merging the transcript-based panel to Compustat. We use two complementary empirical strategies. First, we

estimate dynamic responses around adoption using interaction-weighted event studies for staggered treatment timing following Sun and Abraham (2021), which naturally accommodates installation dynamics. Second, we estimate per-year level effects of being in the adopted state using a fixed-effects 2SLS framework, instrumenting firm adoption with the lagged leave-one-out industry-year diffusion share. Throughout, we emphasize three outcome blocks: (i) pricing power versus cost structure, (ii) intangible deepening and within-firm factor-share reallocation, and (iii) valuation and timing along the industry diffusion curve.

Three patterns emerge. First, production-based markups are flat—if anything mildly declining—around adoption, while cost shares and accounting margins display dynamics consistent with installation and reorganization rather than rising pricing power. Second, adoption is intangible-intensive and coincides with a within-firm reallocation of value added away from labor and toward capital and intangibles. Third, valuation gains are largest for early adopters and decline sharply with industry diffusion at the time of adoption.

Finding 1: AI adoption reorganizes operating costs more than markups. Measured markups are flat around adoption, while operating costs and profitability evolve in a pattern consistent with an installation/build-out phase followed by operating reorganization. Cost shares move even without rising markups, consistent with quasi-fixed overhead and intangible accumulation changing overhead intensity, and the data also indicate that adoption can raise certain variable inputs (e.g., compute and vendor services). Thus, cost reallocation—rather than universal cost decline or increased pricing power—is the empirically relevant margin.

Finding 2: AI is intangible-intensive and shifts value added away from labor. Following adoption, firms increase intangible intensity and intangible capital, and factor shares shift within the firm: the capital share of value added rises while labor's share falls. In the event studies, by year $k = 4$ after adoption the labor share declines by about 5 percentage points, consistent with AI operating as an intangible-heavy technology that

changes organization and production.

Finding 3: the market rewards early adoption; the premium declines with diffusion. Valuation responses to adoption are sharply decreasing in industry diffusion at the time of adoption. Early adopters in low-diffusion environments experience positive revaluations consistent with markets pricing AI as a growth option and rewarding perceived leadership, while late adopters receive smaller revaluations that can turn negative, consistent with defensive catch-up and costly transition. In the event studies, Tobin's q_{alt} reaches about $+0.47$ by year $k = 5$ for first movers versus about $+0.29$ for follower adopters. This diffusion-dependent valuation pattern is reminiscent of earlier general-purpose technology episodes in which market values reallocate toward firms perceived as best positioned to exploit a new technology early.

Contribution and relation to the literature.

A large literature views AI as a general-purpose technology whose economic effects depend critically on complementary investments, organizational change, and diffusion rather than on automation alone (Brynjolfsson and McAfee, 2014; Acemoglu and Restrepo, 2018a). Because effective deployment requires data infrastructure, software, and firm-specific know-how, much of the investment is intangible and imperfectly captured in standard accounts, creating adjustment lags and measurement challenges for productivity and profitability (Brynjolfsson et al., 2017; Corrado et al., 2022).

Recent empirical work has begun to measure firm-level AI adoption and relate it to growth and valuation. Using labor-market signals such as resumes and job postings, Babina et al. (2024) show that AI-investing firms grow faster and experience higher market valuations, largely through increased product innovation and expansion by large incumbent firms. Related work documents that AI investment is associated with organizational restructuring, changes in workforce composition, and flatter hierarchies, consistent with substantial reorganization costs and intangible accumulation (Babina et al., 2023). A limitation of much of this work is that common proxies—labor-market signals, patents, or keyword-based disclosure measures—do not cleanly identify when a firm transitions from

generic AI discussion to operational deployment, nor do they isolate deployment from customer adoption or macro "AI talk"(Mishra et al., 2022; Eisfeldt et al., 2023a).

A related challenge in the literature is measurement. Existing approaches proxy for firms' AI exposure using task-based indices (e.g., O*NET mappings), labor-market signals (job postings and resumes), patent-based measures, or keyword-based textual disclosures. Task-based exposure is useful for characterizing potential labor-market effects but does not identify when a firm actually deploys AI in production (Eloundou et al., 2023). Patent-based measures capture invention rather than implementation and may miss adoption that relies on standard architectures applied to proprietary data and workflows. Labor-market signals primarily reflect capability building that can precede or follow deployment and may understate adoption through third-party vendors and cloud platforms (Babina et al., 2024, 2023). Keyword-frequency disclosure measures are scalable but often conflate general discussion with concrete firm-level use (Mishra et al., 2022). Together, these limitations make it difficult to precisely date adoption events or to separate early deployment from late, defensive catch-up.

This paper contributes by directly dating concrete deployment events from earnings and investor calls using an LLM extraction schema that is pinned to the Capital IQ header firm and returns supporting text for auditability. Linking these deployment dates to Compustat, the analysis separates pricing-power channels from operating-cost reorganization by combining production-based markups with detailed cost shares. This matters because production-based markups are sensitive to production-function assumptions and input-share measurement and need not mechanically rise when firms adopt new technologies (De Loecker and Warzynski, 2012; Benkard et al., 2025). The paper also connects adoption to within-firm factor-share reallocation in the spirit of task-based and automation frameworks that predict shifts toward capital and intangible inputs and away from labor (Autor et al., 2003; Acemoglu and Autor, 2011; Acemoglu and Restrepo, 2018b, 2020).

Finally, the paper links real-side reorganization to market pricing along the diffusion curve. A classic insight from the diffusion literature is that private returns depend on

adoption timing and competitive diffusion: early adopters may capture quasi-rents and option value, while late adopters adopt defensively once technologies become industry standards (Corrado et al., 2021; Berlingieri et al., 2024). In finance, recent work shows that equity markets respond sharply to generative-AI-related news (Eisfeldt et al., 2023b), but it remains unclear whether these valuation effects reflect durable rents, growth options, or diffusion-dependent repricing. By dating deployment at the firm level and interacting adoption with industry diffusion, this paper provides direct evidence on how valuation payoffs erode as AI becomes widespread.

Section 2 describes the transcript corpus, the LLM extraction pipeline, and the construction of firm-year AI adoption measures, including our definitions of adopters, never adopters, and industry diffusion shares. Section 3 lays out the empirical designs– staggered-adoption event studies and an instrumental-variables framework– and specifies the estimation equations and identification strategy. Section 4 reports baseline estimates and heterogeneity by adoption timing and diffusion, and then studies heterogeneity across AI sourcing and the level of integration. Section 5 discusses mechanisms and external validity. Section 6 concludes. The Appendix contains supplementary material and additional robustness checks.

2 Data and Construction

2.1 Transcript universe

We assemble a panel of English, earnings-related transcripts from S&P Capital IQ Transcripts spanning 2011–2024. After de-duplication at the vendor transcript identifier, the corpus contains 187,909 firm–event observations for U.S. public companies.[2] Each transcript is parsed at the firm–event level using Capital IQ's header Target Company identifier. When transcripts discuss multiple firms (e.g., suppliers or peers), we extract information only for the header firm by design. Each observation corresponds to one

[2]Private and non-U.S. companies are excluded; interpret the panel as representative of U.S. public issuers covered by Capital IQ.

Capital IQ transcript event for the header firm, timestamped at the event date/time in the vendor metadata. Figure 1 summarizes the yearly evolution of the transcript universe and the AI-related subsets used in the analysis. Appendix Tables 4 and 5 report transcript counts by year and event-type composition in the full universe.

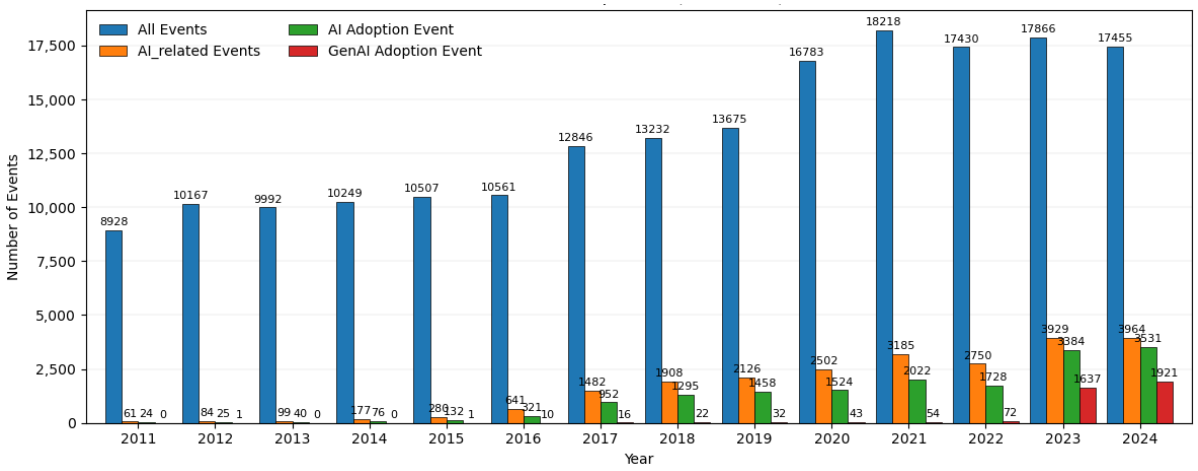


Figure 1: Event counts per year. Blue bars: all transcripts in the universe. Orange bars: transcripts containing AI-related keywords (screened subset). Green bars: events with an explicit AI adoption or exploration signal ($AID \geq 0.5$). Red bars: events with an explicit GenAI adoption or exploration signal ($GenAID \geq 0.5$). Adoption events are defined by the scoring rules in Section 2.3.

2.2 AI keyword screen and analysis subsets

To route potentially AI-relevant transcripts into structured parsing, we apply a deterministic transcript-level keyword screen. The screen retains 23,260 events ($\approx 12.4\%$ of the transcript universe); see the orange series in Figure 1. The keyword screen is used only as an initial routing filter: all retained transcripts are subsequently parsed line-by-line by the LLM, which removes false positives. The only potential omission arises from AI discussions that avoid the screened terms (false negatives). The full keyword dictionary and normalization rules are provided in Appendix A.

From the screened subset, the LLM assigns event-level deployment scores for general AI and GenAI (defined below). We classify an event as containing AI (GenAI) adoption or exploration content if $AID \geq 0.5$ ($GenAID \geq 0.5$). This yields 16,631 events with $AID \geq 0.5$ and 3,821 events with $GenAID \geq 0.5$, corresponding to the green and red

series in Figure 1 and forming the basis for our descriptive evidence and econometric designs.

2.3 LLM extraction and adoption scoring

We transform AI-screened transcripts into structured firm–event observations using one prompt per (transcript, company). The prompt explicitly constrains all answers to the Capital IQ header target company and returns a JSON object that populates the analysis variables.

Definition (AID). For each transcript event e of firm i, the pipeline outputs an ordinal deployment score for (i) general AI, $AID_{i,e} \in \{0, 0.5, 1\}$, and (ii) generative AI, $GenAID_{i,e} \in \{0, 0.5, 1\}$. Scores are defined from transcript evidence as described below; higher values indicate more concrete deployment/integration, and the prompt returns supporting quotes for auditability.

Adoption scoring. For both general AI and GenAI, we score adoption on {0, 0.5, 1}. A score of 1 requires an explicit statement consistent with adoption or deployment (e.g., "we have integrated AI into ..."); 0.5 captures exploratory, pilot, or aspirational language (e.g., "we are exploring AI"); 0 indicates no adoption content. The prompt also returns supporting sentence(s) from the transcript that justify the score. Illustrative examples and the complete prompt schema are reported in Appendix A (see Appendix Figure 14). In the main analysis, we date first adoption using the strict threshold ($AID = 1$ for AI; $GenAID = 1$ for GenAI); we use the lenient threshold ($AID \geq 0.5$; $GenAID \geq 0.5$) for descriptive evidence and robustness. Our adoption date is the first public disclosure consistent with deployment or integration, which may lag internal experimentation; we therefore interpret event-time dynamics as responses around disclosure/implementation rather than the literal start date of internal development. Accordingly, our treatment is best interpreted as the first publicly evidenced transition to deployment/integration, not the first internal experiment.

Other extracted fields. The prompt also records (i) role and sourcing (Developer/Adopter/Both; In-house/External/Both; vendor names if disclosed), (ii) use case and scope (core vs. supportive; application types; business units), and (iii) intensity and depth (integration level, investment level, extent of adoption, and model names when mentioned). The full field list and variable mapping are provided in Appendix A.

2.4 Pipeline 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; non-parseable responses are retried with a minimal wrapper. To guard against hallucinations, the prompt includes three placebo questions—Fictional AI Use Case, Secret AI Department, and Quantum AI Investment—that are expected to return "Information not found" when unsupported by the transcript text. We also manually audit a random subset of events by verifying adoption-support quotes and key labels against the underlying transcript text. Extractions were run with Claude 4 Sonnet (2025 family) in JSON-mode between July–August 2025. Additional implementation details and validation diagnostics are provided in the replication materials and Appendix A.

2.5 Linkage and analysis panels

Firm–event extractions are merged to standard firm identifiers (Capital IQ companyid to gvkey when available). We then construct annual firm panels spanning 2011–2024. For each firm–period we compute: (i) AI and GenAI adoption indicators and first-adoption timing derived from event-level $AID_{i,e}$ and $GenAID_{i,e}$, (ii) role and sourcing tags (including vendor names when disclosed), (iii) application and business-unit tags, and (iv) intensity measures (e.g., integration level). These variables are merged to Compustat outcomes and to GICS industry definitions used in the econometric designs.

Coverage. The resulting annual panel spans 4,313 firms over 2011–2024 (60,382 firm–years). Of these, 33,460 firm–years contain at least one transcript (55.41%), as summarized

in Table 1. Figure 15 shows the distribution of transcripts per firm–year.

Table 1: Panel coverage (annual firm–year)

	Firms	Years	Firm–years	With transcripts	Share with transcripts (%)
Annual panel	4,313	14	60,382	33,460	55.41

Definitions of Compustat variables, fill rules, and data hygiene procedures are reported in Appendix C; outcome construction for intangible-adjusted profitability, markups, and factor shares is in Appendix D.

2.6 Adoption dynamics

Figures 2 and 3 summarize adoption dynamics in the firm-level panel. Unless noted otherwise, these descriptives define adoption using the lenient event-level threshold ($AID \geq 0.5$ for AI; $GenAID \geq 0.5$ for GenAI), aggregated to the firm–year level as described below. For general AI, the share of firms that have ever adopted by year t (stock) rises gradually from near zero prior to 2016 to roughly 31% by 2024, while the share adopting or exploring AI in a given year (prevalence) reaches about 19%. GenAI adoption is much more temporally concentrated: incidence jumps in 2023, and by 2024 about 11% of firms have ever adopted and roughly 9% adopt or explore GenAI in that year. These patterns are consistent with broad diffusion of traditional AI through the late 2010s and a sharp GenAI step-up after 2022.

Outcomes and constructed variables. We harmonize extracted measures into analysis-ready indicators and intensity variables (adoption, sourcing, application tags, business-unit tags, and integration/investment levels). Binary indicators are aggregated to the firm–year level using a logical-any rule, while graded measures are aggregated by mean and sum. Appendix C provides the variable map. Construction of intangible-capital-adjusted profitability, markups, and factor shares follows Peters and Taylor (2017) and De Loecker and Warzynski (2012), with full implementation details in Appendix D.

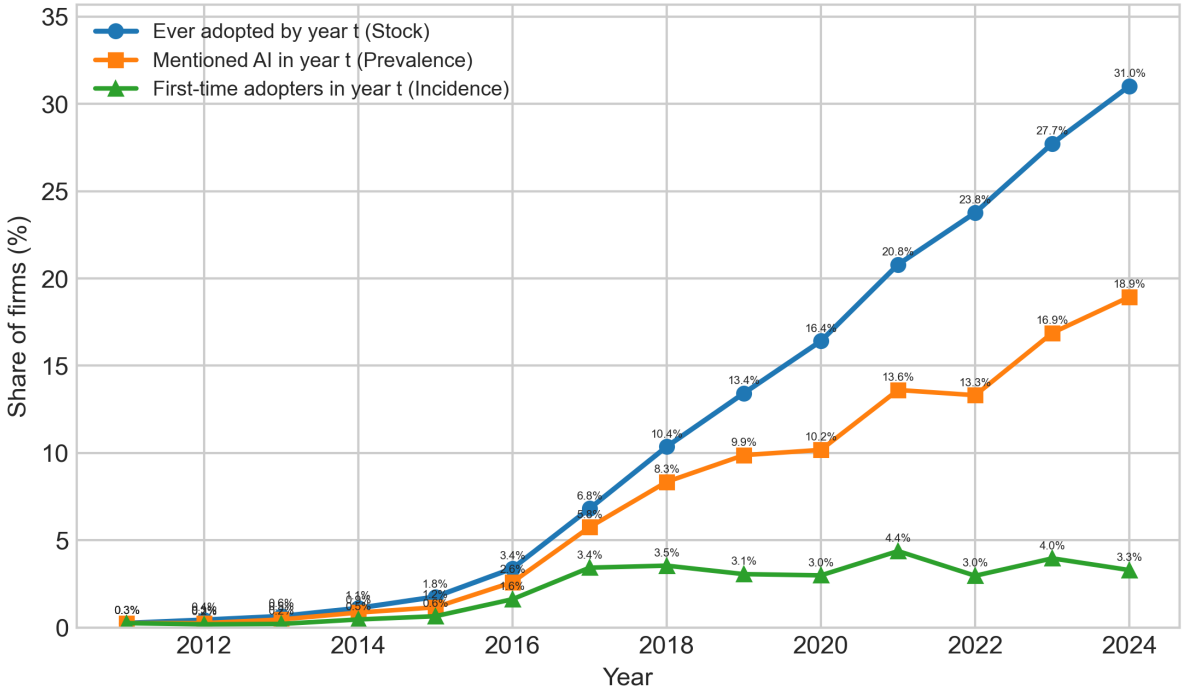


Figure 2: AI adoption dynamics at the firm level. "Stock": share of firms ever adopted by year t. "Prevalence": share with at least one AI-scored transcript event in year t (lenient threshold, $AID \geq 0.5$). "Incidence": first-time adopters in year t (based on firm-level first-adoption timing).

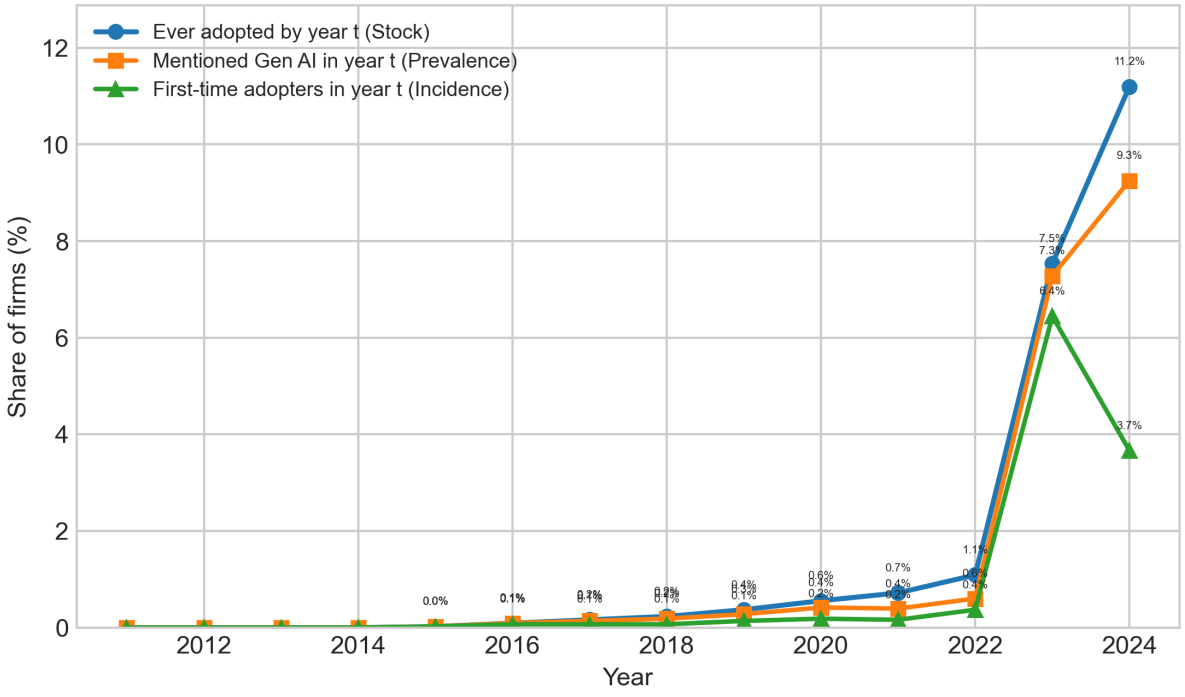


Figure 3: GenAI adoption dynamics at the firm level; definitions as in Figure 2 using the GenAI score (GenAID).

3 Methodology

Our empirical analysis proceeds in three stages that correspond to (i) who adopts and when (selection), (ii) how outcomes evolve around adoption (dynamic treatment effects), and (iii) complementary level evidence on the adopted state from a fixed-effects 2SLS design. Throughout, adoption is treated as absorbing: once a firm is classified as a strict adopter, it remains in the treated state in all subsequent years.

3.1 Notation, treatment definitions, and inference

Let i index firms, t years, and $g(i)$ the firm's GICS industry. Let G_i denote firm i's first strict adoption year. Define event time $K_{it} = t - G_i$ for adopters and the absorbing adoption-state indicator

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

Our strict adoption indicator is binary. We impose absorbing adoption: once a firm is classified as an adopter in year t, it remains an adopter in all subsequent years.

Outcomes are Y_{it}, firm fixed effects are α_i, and year fixed effects are λ_t. Unless otherwise noted, all continuous variables used in the Sun–Abraham event studies and 2SLS specifications are winsorized at the 1st/99th percentiles within year (Appendix C); indicator variables are not transformed. Standard errors are clustered by firm in the event-study specifications. In the 2SLS specifications, baseline standard errors are clustered by firm; in robustness, we also report two-way clustering by firm and year. First-stage diagnostics (Angrist-Pischke partial F and partial R^2) are computed using firm clustering.

Never adopters. We define a never adopter as a firm for which we never observe strict AI adoption in any available transcript through 2024, conditional on having at least one transcript in the sample period. Because all firms in our analysis sample have transcript coverage by construction, this classification reflects the absence of an adoption signal in the firm's transcript record rather than missing disclosure opportunities. This is an observational definition: it identifies firms that never disclose (or otherwise trigger) strict

adoption in the transcript corpus through 2024.

3.2 Selection into adoption: hazard and prediction

Because AI deployment is not randomly assigned, we document selection patterns to discipline interpretation of the event-time and IV results. We summarize selection along three margins: (i) when firms adopt (a discrete-time hazard on a firm–year risk set), (ii) which diffusion-era cohort adopters belong to (early vs. late, among adopters), and (iii) whether firms ever adopt (ever vs. never, cross section). These analyses are descriptive and predictive; they are not used for identification. Their role is to characterize observables associated with adoption and to motivate within-firm designs and diffusion-heterogeneity analyses in the main results.

Modeling approach. Across exercises, our baseline classifier is an L_2-penalized logit (ridge), which accommodates high-dimensional fixed effects and mitigates quasi-separation. We evaluate out-of-sample discrimination using AUC under stratified holdout splits and stratified K-fold cross-validation, and we report odds ratios for standardized continuous covariates as directional summaries under shrinkage. We also estimate a nonlinear classifier (gradient-boosted trees, XGBoost) on the same feature set as a robustness check; it yields similar performance for the hazard and early-vs-late exercises. For the ever-vs-never cross section, however, the nonlinear model materially improves fit (AUC ≈ 0.97), so we emphasize that specification and summarize global feature importance using SHAP in the Appendix.

3.2.1 Discrete-time adoption hazard

Let G_i denote firm i's first strict AI deployment year (as defined in Section 2.3, using $AID = 1$), and define the first-adoption event

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

We construct the discrete-time risk set as all firm–years up to and including first adoption for eventual adopters ($t \le G_i$) and all firm–years for never adopters. The baseline hazard is an L_2-penalized logit with year and industry fixed effects:

$$\Pr\big(\text{Adopt}_{it} = 1 \mid t \le G_i \text{ or } G_i \text{ undefined}\big) = \Lambda\Big(\lambda_t + \gamma_{g(i)} + \boldsymbol{\beta}' X_{it}\Big), \tag{3.1}$$

where $\Lambda(\cdot)$ is the logistic CDF, λ_t are year effects, $\gamma_{g(i)}$ are GICS-industry effects, and X_{it} collects firm fundamentals measured in the risk set:

$$\begin{aligned} X_{it} = \{&\ln(\text{mktcap}_{it}),\ R\&D/AT_{it},\ q_{it},\ \text{intangibles}_{it},\ \text{book leverage}_{it},\ \text{Capex/AT}_{it},\\ &\text{Cash/AT}_{it},\ \text{operating margin}_{it},\ \text{disclosure score}_{it},\ S^{\text{AI}}_{g(i),t-1}\}. \end{aligned} \tag{3.2}$$

The diffusion covariate $S^{\text{AI}}_{g(i),t-1}$ is the lagged leave-one-out industry adoption share, included to capture diffusion pressure faced by firms in the same industry.[3] Continuous covariates are median-imputed and standardized prior to estimation. We interpret the fitted model as a descriptive summary of observable predictors of adoption timing, not as a structural adoption model.

3.2.2 Vintage selection among adopters: early vs. late cohorts

To connect selection patterns to the diffusion heterogeneity studied in the main results, we compare early vs. late adopters within the set of eventual adopters. Define

$$\text{Early}_i = \mathbf{1}\{G_i \le 2018\}, \qquad \text{Late}_i = \mathbf{1}\{G_i \ge 2019\}.$$

We collapse the panel to one baseline row per adopting firm using pre-adoption data whenever possible: we take the median of fundamentals over $t \le G_i - 1$ (falling back to the earliest available year if no pre-adoption observations exist), and define baseline industry as the modal GICS classification over the same window. We then estimate a

[3] This is the same object used for the diffusion-based instrument, lagged and leave-one-out by construction.

ridge-logit classifier:

$$\Pr(\text{Early}_i = 1) = \Lambda\Big(\kappa + \gamma_{g(i)} + \boldsymbol{\theta}'\bar{X}_i\Big), \tag{3.3}$$

where $\bar{X}_i$ is defined by (3.2) evaluated on the baseline window. To avoid sparse fixed-effect issues, industries with small cell counts are grouped into an "other" category. As above, continuous features are median-imputed and standardized, and degenerate predictors are dropped.

3.2.3 Ever vs. never adoption

Finally, we summarize cross-sectional differences between firms that ever adopt and firms that never adopt. Let $\text{EverAdopt}_i = 1$ if G_i is defined and $\text{EverAdopt}_i = 0$ otherwise. We construct one baseline row per firm: for adopters we use pre-adoption years $t \leq G_i - 1$ when available (otherwise the earliest year), while for never adopters we use the earliest year in sample. In addition to industry indicators, we include a baseline-year fixed effect (the first year of the baseline window) to absorb differences in sample entry timing.

We estimate both ridge logit and XGBoost on the same baseline feature set. In this ever-vs-never exercise, the nonlinear model achieves very high discrimination (AUC ≈ 0.97), so we report it and summarize global feature contributions using SHAP in the Appendix. We keep this cross section secondary in the main text because the hazard specification embeds never adopters in a timing-consistent risk set; the cross section is included mainly as a compact "who adopts" profile and a robustness check.

Across exercises, the takeaway is descriptive: adoption timing and adopter composition are strongly predictable from pre-adoption fundamentals, industry exposure, and diffusion pressure. This motivates our emphasis on within-firm event-time designs and on diffusion-heterogeneity analyses, rather than treated-versus-untreated comparisons.

3.3 Dynamic treatment effects: Sun–Abraham event studies

We estimate dynamic effects around adoption under staggered timing using the interaction-weighted (IW) event-study estimator of Sun and Abraham (2021). This approach is robust

to treatment-effect heterogeneity across cohorts and event time and avoids the negative-weight problem that can arise in two-way fixed-effects event studies. In all event-study specifications, "adoption" is defined using the strict deployment threshold (first year with $AID = 1$ unless stated otherwise).

Setup. Let G_i denote firm i's first adoption year (set $G_i = \infty$ for never adopters) and define event time $K_{it} = t - G_i$. We estimate cohort-by-event-time effects using firm and year fixed effects, where for each cohort g and relative time k the comparison group consists of not-yet-treated firms (and never-treated firms, when present) observed at the same calendar time.

Cohort $\times$ event-time specification. For each adoption cohort g and event time k, we estimate the saturated interaction specification

$$Y_{it} = \alpha_i + \lambda_t + \sum_{g} \sum_{k \in \mathcal{K} \setminus \{-1\}} \theta_{k,g} \, \mathbf{1}\{G_i = g\} \, \mathbf{1}\{K_{it} = k\} + \varepsilon_{it}, \tag{3.4}$$

where α_i and λ_t are firm and year fixed effects and $k = -1$ is the omitted reference period. Identification relies on (i) parallel trends between treated and not-yet-treated/never-treated firms, (ii) no anticipation, and (iii) sufficient overlap in event-time support.

Interaction-weighted (IW) aggregation. Following Sun and Abraham (2021), we aggregate cohort-specific estimates into an IW event-time path:

$$\beta_k^{\mathrm{IW}} = \sum_{g} \omega_{k,g} \, \theta_{k,g}, \qquad \omega_{k,g} = \frac{\sum_{i:\, G_i = g} \sum_{t} \mathbf{1}\{K_{it} = k\}}{\sum_{g'} \sum_{i:\, G_i = g'} \sum_{t} \mathbf{1}\{K_{it} = k\}}. \tag{3.5}$$

The weights $\omega_{k,g}$ are proportional to the number of cohort-g observations that contribute to event time k in the interacted design. All baseline event-study figures report $\widehat{\beta}_k^{\mathrm{IW}}$ with 95% confidence intervals; standard errors are clustered by firm.

3.3.1 Event windows, binning, support filters, and pretrend tests

We use event windows $\mathcal{K} = \{-L, \ldots, -2, 0, \ldots, R\}$ and bin observations outside the window into end bins ($k = -L$ and $k = R$). To avoid reporting mechanically identified coefficients with thin treated support, we impose minimum-support requirements at the event-time level: each reported k must be supported by at least a minimum number of treated firms (and, in subgroup analyses, by a minimum number of cohorts). Event-study figures report the event-time support $N(k)$ (treated firms contributing to k) alongside $\widehat{\beta}_k^{\mathrm{IW}}$.

For each outcome, we report a joint pretrend test of

$$H_0 : \ \beta_k^{\mathrm{IW}} = 0 \quad \forall k < 0,$$

implemented as a Wald test using the firm-clustered covariance matrix.

3.3.2 Early vs. late adopters (diffusion vintages)

Because diffusion accelerates around 2019, we estimate vintage-specific dynamics by splitting adopters into early and late cohorts and re-estimating (3.4) within each subsample:

$$\text{Early adopters: } G_i \leq 2018, \qquad \text{Late adopters: } G_i \geq 2019,$$

always retaining never adopters as controls. Within each vintage we apply the same event-window binning and support filters and compute the IW aggregation (3.5). We report vintage-specific pretrend Wald tests.

3.3.3 First movers vs. followers within industries

We proxy the diffusion environment using a predetermined leave-one-out (LOO) industry adoption share. Let $g(i)$ denote firm i's GICS 6-digit industry and let G_i denote its adoption year. Define

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

the share of other firms in i's industry that have adopted by the year prior to i's adoption. Measuring diffusion at $G_i - 1$ makes the split predetermined at adoption and avoids mechanical same-year feedback.

We classify firms within industry. Among adopters in industry $g(i)$, firm i is a first mover if S_i lies below the 25th percentile of $\{S_j : j \in g(i)\}$; otherwise i is a follower. In robustness, we (i) shift the predetermined diffusion state to $S^{-i}_{g(i),\,G_i-2}$ and (ii) use contemporaneous diffusion $S^{-i}_{g(i),\,G_i}$.

We implement the comparison using separate Sun–Abraham event studies: we estimate (3.4) on (i) first movers plus never adopters and (ii) followers plus never adopters, and compare the implied IW paths $\widehat{\beta}^{\mathrm{IW}}_k$. To ensure comparability, we hold the event window fixed at $k \in [-5, +5]$ in both estimations.

Interpreting diffusion: competitive crowding vs. industry opportunity. The diffusion state $S^{-i}_{g(i),t}$ can be interpreted as a within-industry state variable capturing how common AI deployment has become. A natural mechanism is competitive crowding: when diffusion is low, adoption may be differentiating; when diffusion is high, adoption is closer to a parity technology and incremental rents may be harder to sustain. An alternative interpretation is that diffusion proxies for time-varying industry opportunity or industry-year shocks that jointly move adoption and outcomes; we address this concern by (i) using predetermined diffusion states and (ii) emphasizing within-firm dynamics and robustness across industry definitions.

3.3.4 Controls policy

To preserve the interpretation of event-time coefficients, our baseline event studies include firm and year fixed effects and do not add time-varying controls.

3.4 Instrumental-variables strategy (2SLS)

We complement the Sun–Abraham event-study evidence with a parsimonious level specification that estimates the average effect of being in the adopted state in a given year. The

2SLS design uses lagged within-industry diffusion as a shifter of firms' adoption propensity and is best interpreted as a benchmark for magnitudes under additional assumptions. Throughout, our primary causal evidence remains the event-time dynamics and their pretrend/no-anticipation diagnostics; the IV results are presented as supportive level comparisons and as a check on the direction and scale of the estimated effects.

Treatment states (absorbing adoption). Recall that the event-level deployment measures are the evidence-gated scores $AID_{i,e} \in \{0, 0.5, 1\}$ and $GenAID_{i,e} \in \{0, 0.5, 1\}$ for transcript event e of firm i (Section 2.3). We construct strict annual adoption indicators by setting the firm-year score equal to one if any transcript event in that year attains the strict threshold. Let

$$\widetilde{D}_{it}^{AI} = \mathbf{1}\{\exists e \in (i,t) : AID_{i,e} = 1\}, \qquad \widetilde{D}_{it}^{Gen} = \mathbf{1}\{\exists e \in (i,t) : GenAID_{i,e} = 1\}.$$

We then impose absorbing treatment by cumulating these annual indicators:

$$D_{it}^{AI} = \max_{s \le t} \widetilde{D}_{is}^{AI}, \qquad D_{it}^{Gen} = \max_{s \le t} \widetilde{D}_{is}^{Gen}.$$

Equivalently, $D_{it}^{AI} = \mathbf{1}\{t \ge G_i^{AI}\}$ and $D_{it}^{Gen} = \mathbf{1}\{t \ge G_i^{Gen}\}$, where G_i^{AI} and G_i^{Gen} are the first strict adoption years.

3.4.1 Instrument: lagged leave-one-out industry adoption share

Our baseline instrument is lagged within-industry diffusion, constructed as a leave-one-out (LOO) share. Because adoption is measured from transcript evidence, we form diffusion shares on the set of firms with observable transcript coverage in the relevant industry–year. Let $\mathcal{I}_{g,t}$ denote the set of firms in GICS industry g with at least one transcript in year t, and let $N_{g,t} = |\mathcal{I}_{g,t}|$. For a generic treatment state $D_{it} \in \{D_{it}^{AI}, D_{it}^{Gen}\}$, define

$$Z_{g(i),t-1} = \frac{1}{N_{g(i),t-1} - 1} \sum_{j \in \mathcal{I}_{g(i),t-1},\, j \ne i} D_{j,t-1}. \tag{3.6}$$

Lagging reduces mechanical contemporaneous feedback and aligns the diffusion state with a predetermined shifter of current adoption.

3.4.2 2SLS specification and interpretation

For each outcome Y_{it}, we estimate

$$D_{it} = \pi_0 + \pi_1 Z_{g(i),t-1} + \boldsymbol{\pi}_2^\top \mathbf{X}_{it-1} + \alpha_i + \lambda_t + u_{it}, \tag{3.7}$$

$$Y_{it} = \beta\, D_{it} + \boldsymbol{\gamma}^\top \mathbf{X}_{it-1} + \alpha_i + \lambda_t + \varepsilon_{it}, \qquad D_{it} \text{ instrumented by } Z_{g(i),t-1}, \tag{3.8}$$

where α_i and λ_t are firm and year fixed effects. The baseline control set is intentionally parsimonious:

$$\mathbf{X}_{it-1} = \big\{ \ln(\text{mktcap}_{it-1}),\ (R\&D/AT)_{it-1} \big\},$$

and richer specifications are reported in the Appendix.

We report standard first-stage diagnostics (Angrist–Pischke partial F and partial R^2). Under relevance and monotonicity, β has the usual LATE interpretation for firms whose adoption status is shifted by lagged peer diffusion. The key identifying restriction is exclusion: conditional on firm and year fixed effects and lagged controls, $Z_{g(i),t-1}$ affects Y_{it} only through own adoption. Because diffusion may proxy for industry-level competitive or technological forces that directly affect outcomes, we view the IV estimates as benchmark magnitudes rather than stand-alone identification.

Robustness to industry-level confounds. Because within-industry diffusion may proxy for industry-level shocks that directly affect outcomes, we evaluate sensitivity along three dimensions. First, in addition to the baseline lagged diffusion shifter $Z_{g(i),t-1}$, we re-estimate the 2SLS using the contemporaneous diffusion share $Z_{g(i),t}$ as an alternative (less predetermined) instrument to assess whether results are driven mechanically by the one-year lag choice. Second, we augment the baseline specification with lagged industry–year averages of slow-moving proxies for industry conditions (e.g., sales growth, operating margin, and R&D intensity), constructed at the industry–year level and merged back to

the firm panel. Third, we reconstruct the diffusion shifter at a coarser industry aggregation (e.g., sector-level) and re-estimate the vintage-interacted 2SLS to reduce simultaneity at very granular industry definitions. Across these variations, the estimated magnitudes are similar. Detailed results are reported in Appendix Tables 16, 17 and 18.

3.4.3 Vintage-heterogeneous 2SLS (early vs. late adopters)

Motivated by the diffusion break around 2019, we allow the level effect of adoption to differ across adoption vintages. Let $G_i \equiv G_i^{AI}$ and define

$$D_i^{\text{early}} = \mathbf{1}\{G_i < 2019\}, \qquad D_i^{\text{late}} = \mathbf{1}\{G_i \geq 2019\},$$

with $D_i^{\text{early}} = D_i^{\text{late}} = 0$ for never adopters. We form two endogenous regressors by interacting the adoption state with vintage:

$$D_{it}^{AI,early} = D_{it}^{AI} \cdot D_i^{\text{early}}, \qquad D_{it}^{AI,late} = D_{it}^{AI} \cdot D_i^{\text{late}},$$

and instrument them using the corresponding interacted diffusion shifters:

$$Z_{g(i),t-1}^{AI,early} = Z_{g(i),t-1}^{AI} \cdot D_i^{\text{early}}, \qquad Z_{g(i),t-1}^{AI,late} = Z_{g(i),t-1}^{AI} \cdot D_i^{\text{late}}.$$

The second stage is

$$Y_{it} = \beta_{\text{early}} D_{it}^{AI,early} + \beta_{\text{late}} D_{it}^{AI,late} + \boldsymbol{\gamma}^\top \mathbf{X}_{it-1} + \alpha_i + \lambda_t + \varepsilon_{it}, \tag{3.9}$$

instrumenting $\{D_{it}^{AI,early}, D_{it}^{AI,late}\}$ with $\{Z_{g(i),t-1}^{AI,early}, Z_{g(i),t-1}^{AI,late}\}$. We report β_{early} and β_{late} and test $H_0 : \beta_{\text{early}} = \beta_{\text{late}}$. This split is naturally interpreted as heterogeneity in level effects by realized adoption vintage; it is most informative when the diffusion shifter changes the probability of adoption within vintage rather than moving firms across the 2019 boundary.

3.4.4 Implementation details

All 2SLS specifications include firm and year fixed effects and the lagged controls $\mathbf{X}_{it-1}$. Baseline standard errors are clustered by firm. Appendix robustness specifications are reported with two-way clustering (firm and year); see Tables 16–18.

3.5 Heterogeneity and mechanisms

We present two complementary extensions of the baseline event-study design. First, we study heterogeneity by adoption mode using group-specific dynamic effects within classified-only subsamples. Second, we study mechanisms by relating dynamic effects to a time-varying, standardized measure of AI integration intensity and plotting predicted paths at low vs high intensity.

Heterogeneity by adoption mode (role and sourcing). To examine whether the dynamic effects of AI adoption vary by adoption mode, we estimate heterogeneous Sun–Abraham event studies that allow treatment effects to differ across a binary indicator H_i capturing (i) external vs. In-house sourcing and (ii) role classification (user or vendor). For each heterogeneity dimension we restrict attention to a classified-only sample, so that $H_i = 0$ denotes the relevant comparison group within the classified subset rather than firms with missing/unknown classification. We then interact event-time indicators with H_i, producing group-specific dynamic effects $\{\beta_k^{H=0}, \beta_k^{H=1}\}_k$ relative to the omitted pre-adoption baseline period. The figures report these two paths along with the event-time sample sizes $N(k)$.

Mechanisms: event-time effects by integration intensity. To probe mechanisms, we relate post-adoption dynamics to the time-varying intensity of AI integration. Let Z_{it} denote a standardized measure of integration intensity observed at the firm-year level (e.g., cumulative integration in logs or integration level). We estimate a Sun–Abraham event-study specification augmented with interactions between event-time indicators and

Z_{it}, allowing the dynamic effect at horizon k to vary with integration intensity:

$$Y_{it} = \cdots + \sum_{k \neq k_0} (\beta_k + \gamma_k Z_{it}) \, \mathbf{1}\{K_{it} = k\} + \varepsilon_{it},$$

where K_{it} is event time and k_0 is the omitted baseline pre-adoption period. We then report predicted event-time effects evaluated at $Z = 0$ (average intensity) and $Z = +1$ (one standard deviation above average). Because Z_{it} is time-varying and may itself respond to post-adoption decisions, these results are interpreted as dose–response/channel evidence rather than causal heterogeneity by predetermined firm type. Therefore, since Z_{it} evolves after adoption, the intensity interactions should be viewed as descriptive evidence on channels (how outcomes covary with deeper integration), not as clean exogenous variation in treatment intensity. Appendix Figure 18 summarizes the time-series distribution of the disclosure-based AI integration level used in this section.

4 Results

This section presents three sets of results that map to our empirical mechanism and interpretation (Section 5): (i) who adopts AI and when, (ii) dynamic responses of costs, factor shares, and valuation around adoption, and (iii) level (per-year) effects from a complementary fixed-effects IV design. Throughout, interpret the adoption year G_i as the first observable disclosure consistent with deployment/integration in the firm's investor communications, not necessarily the internal start date of experimentation.

4.1 Who adopts AI, and when do they adopt?

We begin by describing (i) which firms adopt AI and (ii) when adoption occurs along the industry diffusion curve. This evidence serves two purposes. First, it shows that adoption is highly systematic rather than mechanical or random: it is strongly predicted by firm fundamentals and the diffusion state of the industry. Second, it motivates the diffusion- and timing-based partitions used below (early vs. late; first movers vs. followers) and

anticipates the timing heterogeneity in valuation and cost structure documented below.

4.1.1 Adoption hazard: diffusion and firm fundamentals

Design. We estimate the discrete-time adoption hazard and the auxiliary prediction exercises described in Section 3.2. Briefly, the baseline specification is a ridge-logit hazard with year and industry fixed effects, and we report odds ratios for standardized covariates; nonlinear classifiers (XGBoost) and SHAP summaries are used as descriptive diagnostics (Appendix E).

Magnitudes. Table 2 reports odds ratios for standardized covariates. Two patterns emerge. First, industry diffusion is the dominant correlate of adoption: a one-standard-deviation increase in the leave-one-out lagged industry adoption share $S^{-i}_{g,t-1}$ is associated with roughly a doubling of adoption odds in the hazard model (OR $\approx$ 2.01). Second, adopters are larger, more cash-rich, and more intangible-intensive: log(mktcap), cash/assets, and intangibles/assets load positively, while book leverage loads weakly negatively. Predictive performance is high (ridge-logit CV mean AUC 0.824 with sd 0.013), indicating that adoption is strongly predictable from observables.

Table 2: Predictors of AI adoption: hazard, ever-vs-never, and early-vs-late (ridge logits)

	Hazard OR	**Ever vs. Never OR**	**Early vs. Late OR**
Industry AI diffusion $S^{-i}_{g,t-1}$ (LOO, lag 1)	2.006	3.951	—
log(Market Cap)	1.413	2.284	1.189
Cash / Assets	1.423	1.629	1.116
Intangibles / Assets	1.338	1.457	1.135
Operating margin	1.092	1.077	2.290
Disclosure score	1.078	0.888	0.990
EV-based q proxy (q^{EV})	1.067	1.211	0.969
R&D / Assets	1.023	0.747	1.171
Capex / Assets	0.998	0.959	1.291
Debt / Assets (book leverage)	0.968	1.011	0.749

Notes: Odds ratios shown for standardized continuous covariates. Hazard model is a ridge logit with year and industry FE (CV AUC 0.824 ± 0.013). Ever-vs-never is a firm-level ridge logit using pre-adoption baselines for adopters and all pre-2025 observations for never-adopters (CV mean AUC 0.857 with sd 0.012; XGBoost CV mean AUC 0.981 with sd 0.003). Early-vs-late is estimated on adopters only (early = 2011–2018, late = 2019–2024; in-sample AUC 0.711). "—" indicates the diffusion regressor is not included in the early-vs-late classification.

4.1.2 Timing within adopters: early vs. late adoption

Within the set of eventual adopters, early adoption is concentrated among higher-performance firms. In Table 2, the early-vs-late odds ratios show that early adopters are substantially more profitable (operating margin OR ≈ 2.29) and are more capital- and innovation-intensive conditional on other covariates (Capex/Assets OR ≈ 1.29; R&D/Assets OR ≈ 1.17), while carrying lower book leverage (Debt/Assets OR ≈ 0.75). This selection pattern is useful for interpreting the diffusion-dependent valuation results below: early adopters look like higher-capability firms with stronger complementary assets, whereas late adopters appear closer to defensive catch-up adopters operating in a mature diffusion state.

Appendix diagnostics. Appendix E reports SHAP summaries for the hazard and ever-vs-never classifiers (Figures 34–35) and the industry fixed-effect odds ratios for the hazard and ever-vs-never exercises (Tables 12–13).

4.2 Dynamic effects around adoption

We now study dynamic adjustment paths using interaction-weighted event studies for staggered treatment timing (Sun–Abraham; Section 3.3). Figures report event-time coefficients $\hat{\beta}_k$ for $k = -5, \ldots, 5$ (relative to the omitted reference year $k = -1$). Shaded bands denote 95% confidence intervals for the aggregate IW estimates, and markers report the number of treated observations contributing at each horizon, $N(k)$.

Across outcomes, the dominant patterns are: (i) operating cost shares and accounting margins display an installation/retooling transition, (ii) markups do not exhibit a sustained post-adoption increase, (iii) intangible intensity and intangible capital rise, (iv) factor shares shift away from labor, and (v) valuation responses are sharply heterogeneous by adoption timing and diffusion state.

4.2.1 Scale and operating cost shares

Figure 4 reports dynamics in sales levels and the two key operating cost ratios, COGS/Sales and SG&A/Sales.

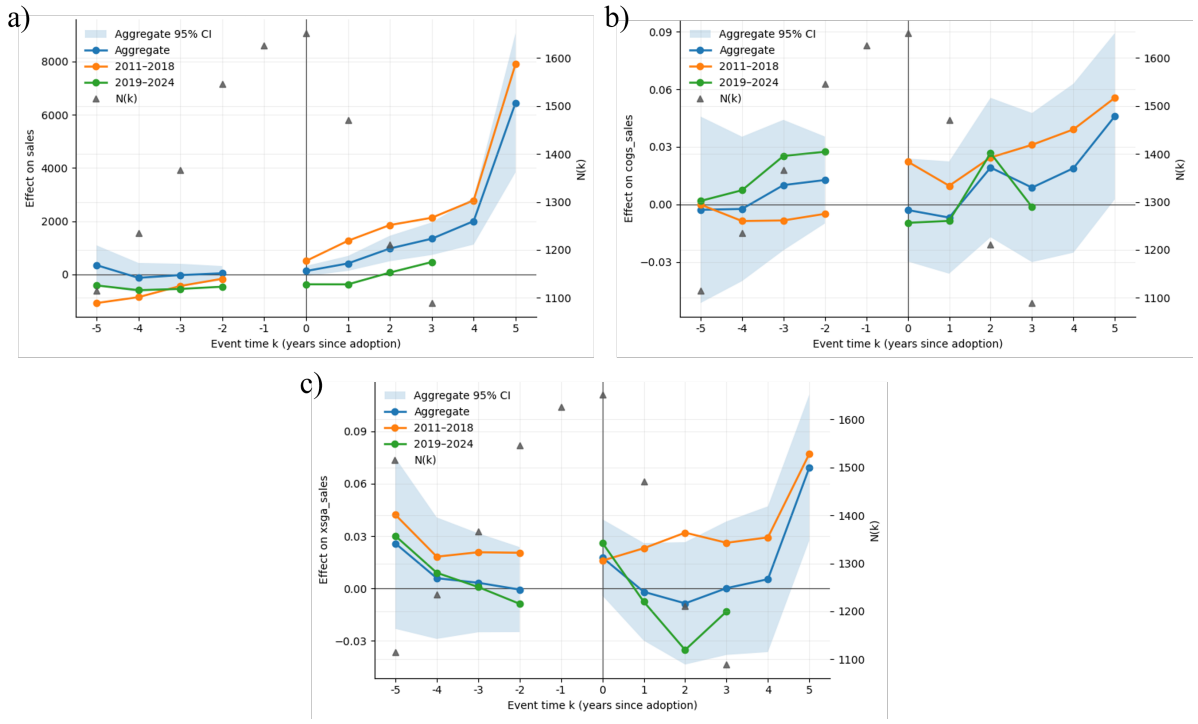


Figure 4: Sun–Abraham IW event studies for scale and cost shares: Sales (level), COGS/Sales, and SG&A/Sales. The aggregate IW path is shown with 95% confidence bands; overlaid paths split adopters by vintage (early 2011–2018 vs. late 2019–2024).

Sales (scale expansion). Sales are approximately flat in the pre-period and then rise after adoption. The aggregate path shows a gradual post-adoption ramp that becomes economically large by $k = 4$–5. The vintage overlay highlights that scale expansion is concentrated among early adopters: the 2011–2018 cohort exhibits much larger sales gains by $k = 5$ than the 2019–2024 cohort, whose post-adoption sales response is smaller within the shorter support window.

COGS/Sales (variable cost intensity). COGS/Sales does not fall systematically after adoption. Instead, the aggregate series trends upward in the post-period, indicating that AI adoption is compatible with higher measured variable costs per dollar of revenue. This pattern is consistent with recurring AI usage inputs recorded in COGS (e.g., compute, data, model-serving costs, vendor fees proportional to usage). Across vintages, both early and late adopters exhibit upward pressure on COGS/Sales, though the timing differs and later-horizon estimates are noisier.

SG&A/Sales (overhead and build-out). SG&A/Sales exhibits the clearest timing heterogeneity. Early adopters show elevated SG&A/Sales around and after adoption (consistent with expensed build-out and organizational redesign), whereas late adopters display a decline in SG&A/Sales in the post period over the common-support window. This pattern is consistent with a frontier-versus-catch-up interpretation: when diffusion is low, adoption is more experimental and build-out intensive, whereas later adoption in high-diffusion environments is more standardized.

Where do AI-related costs appear in financial statements? COGS vs. SG&A. A natural concern is that AI adoption should mainly raise SG&A rather than cost of goods sold. Conceptually, AI-related costs can enter operating expenses through at least two channels. First, product-embedded AI (e.g., inference for customer-facing features, personalization, content generation, fraud detection at transaction scale) can raise cost of revenue and therefore appear in COGS. Second, internal deployment costs (reorganization, process automation, model development and monitoring, data engineering, cloud subscriptions, vendor contracts, compliance and governance) are more naturally recorded in SG&A. In addition, adoption may induce reclassification and product-mix changes, so the COGS versus SG&A split is ultimately an empirical question.

We therefore complement the aggregate event-study evidence with industry splits that separate settings where product-embedded AI is more plausible from settings dominated by internal automation. The patterns are consistent with this mapping: in Information Technology, post-adoption effects are positive for both COGS/Sales and SG&A/Sales, while in Financials we observe a pronounced rise in SG&A/Sales with little or no corresponding increase in COGS/Sales. Details and full dynamic paths are reported in Appendix M.

4.2.2 Margins and a cost-share decomposition

Figure 5 traces operating profitability measured in two ways: the conventional operating margin and an intangible-adjusted operating margin following Peters and Taylor (2017).[4]

[4]Appendix D.1 details the Peters–Taylor construction: perpetual-inventory stocks for R&D and the organizational-capital component of SG&A, economic amortization, and the adjusted operating-income/margin definitions.

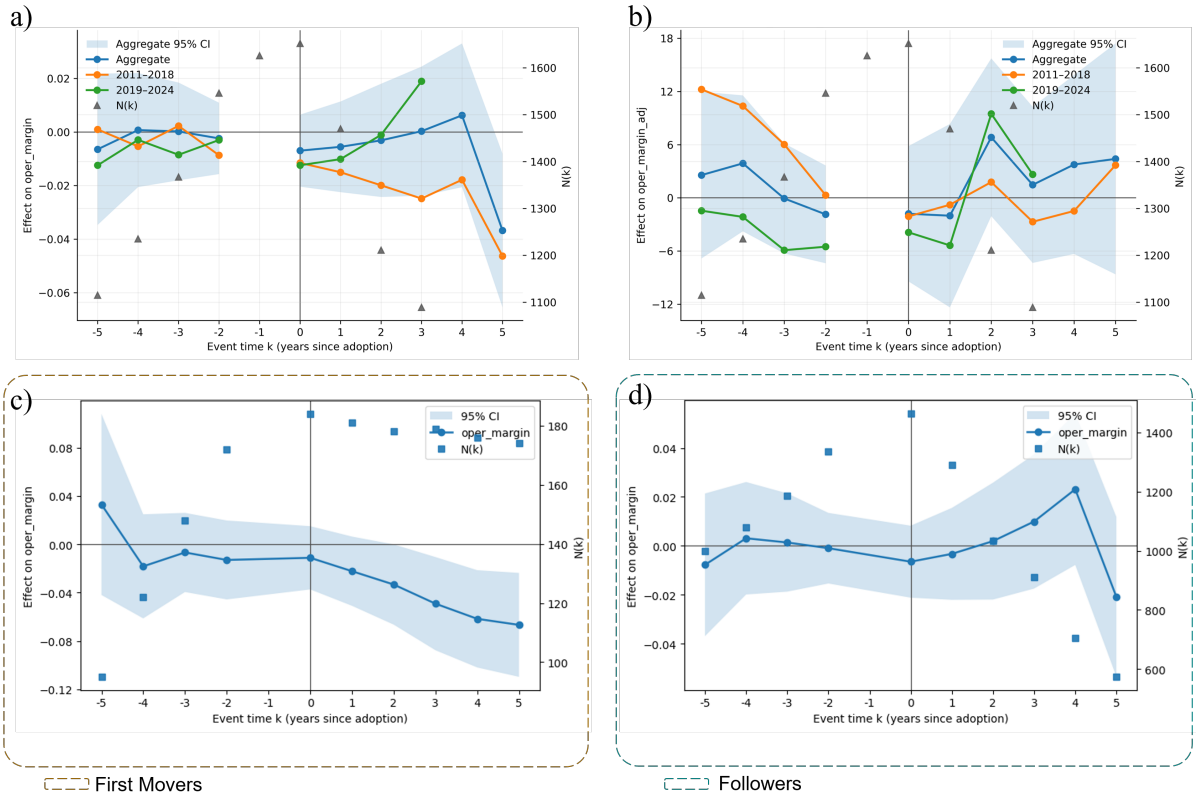


Figure 5: Operating profitability around adoption. Top panels: aggregate Sun–Abraham IW effects on operating margin and an intangible-adjusted operating margin. Overlaid lines split adopters by vintage. Bottom panels: interaction-weight paths for first movers (left) vs. followers (right) (definition in Section 4.3).

Accounting operating margins. Operating margins are close to flat in the aggregate through the medium run and do not show a robust sustained improvement within the five-year window. The vintage split is informative: early adopters exhibit a persistent margin compression after adoption, while late adopters display a milder pattern and, in some horizons, modest improvements. This mirrors the cost-share evidence above: early adopters scale and build organizational capability while expensing a larger share of build-out costs.

Intangible-adjusted profitability. Once we capitalize intangible investment (R&D and the organizational-capital component of SG&A), the post-adoption profitability profile is more positive. In the adjusted series, profitability improves after adoption for both vintages, with especially sharp medium-run movements for late adopters over the common-support window. This mirrors the standard accounting-versus-economic-profit

distinction: expensed intangible investment can depress GAAP margins during installation even if underlying efficiency and expected profitability are improving once intangibles are capitalized (Peters and Taylor, 2017).

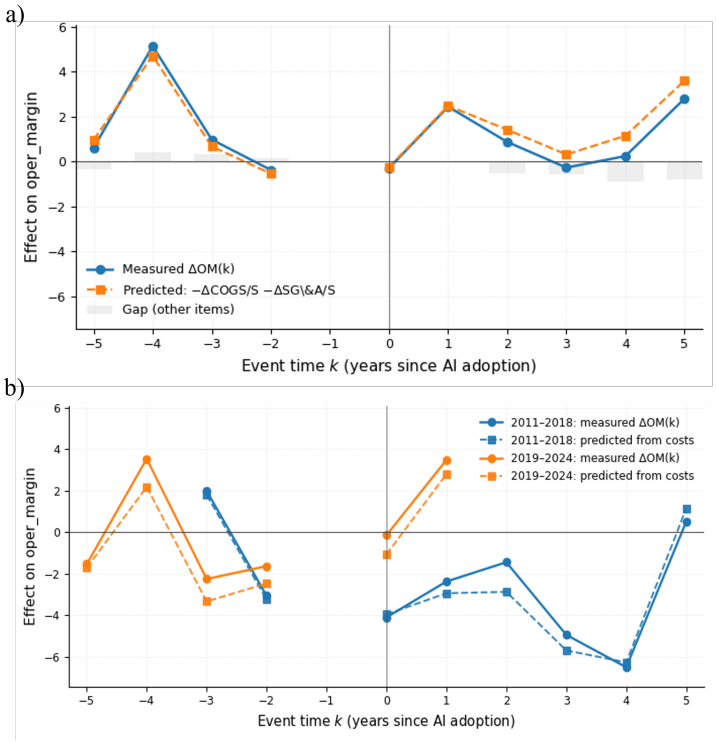


Figure 6: Operating-margin decomposition. The dashed series is the cost-share component $-\Delta(\text{COGS}/S) - \Delta(\text{SG\&A}/S)$; grey bars are the residual Gap(k). The cost-share component closely overlays the measured $\Delta OM(k)$, indicating that operating-margin dynamics are driven by COGS and SG&A shares.

A decomposition: margins move through cost shares. To isolate mechanism, we decompose the change in operating margin at event time k into the contribution of the two major cost ratios and a residual:

$$\Delta OM(k) \;=\; -\,\Delta\big[\text{COGS}/S\big](k) \;-\; \Delta\big[\text{SG\&A}/S\big](k) \;+\; \text{Gap}(k), \tag{4.1}$$

where Gap(k) collects other items (e.g., non-operating components netted out by the accounting definition). Figure 6 shows that the cost-share component tracks $\Delta OM(k)$ closely, and the residual Gap(k) is small. Thus, operating-margin dynamics are overwhelmingly a cost-share story, not a markup story: margins change because COGS/Sales and SG&A/Sales move, not because of a large residual component. The same accounting

identity also explains the vintage split: differences in SG&A/Sales and COGS/Sales account for the divergence in margins between early and late adopters.

4.2.3 Markups and pricing power: no robust post-adoption increase

Production-based markups are sensitive to both production-function assumptions and the measurement of the flexible-input revenue share. We therefore use markups primarily as a diagnostic for whether the profitability and valuation patterns documented above could plausibly reflect rising pricing power.

We study three COGS-share production-based constructions of markups that differ only in how Finance & Insurance firms (NAICS 52) are treated when forming the flexible-input share: (i) a baseline full-sample series, (ii) an "F&I-adjusted" series that augments COGS by interest expense for NAICS 52 firms when interest expense is observed, and (iii) a Finance & Insurance exclusion series that drops NAICS 52 firms entirely. Construction details are provided in Appendix D.2.

Figure 7 reports event-time dynamics in log markups for these three constructions, alongside the gross margin (gross profit over sales). Across constructions, we do not find evidence of a sustained post-adoption increase in pricing-power proxies: aggregate estimates remain close to zero over most horizons and do not display a persistent upward shift in the post-adoption period. If anything, the Finance & Insurance exclusion series shows mild downward drift at longer horizons, reinforcing the interpretation that medium-run profitability dynamics are unlikely to be driven primarily by rising markups.

4.2.4 Intangible deepening and the investment mix

Figure 8 shows the investment side: the intangible asset ratio, R&D/Assets, Capex/Assets, and the level of the intangible capital stock K^{intan}.

Intangibles rise; tangible Capex does not. Intangibles/Assets trends upward after adoption in the aggregate, indicating balance-sheet intangible deepening. By contrast, Capex/Assets is close to flat or mildly declining, providing no evidence of a broad tangible-investment boom.

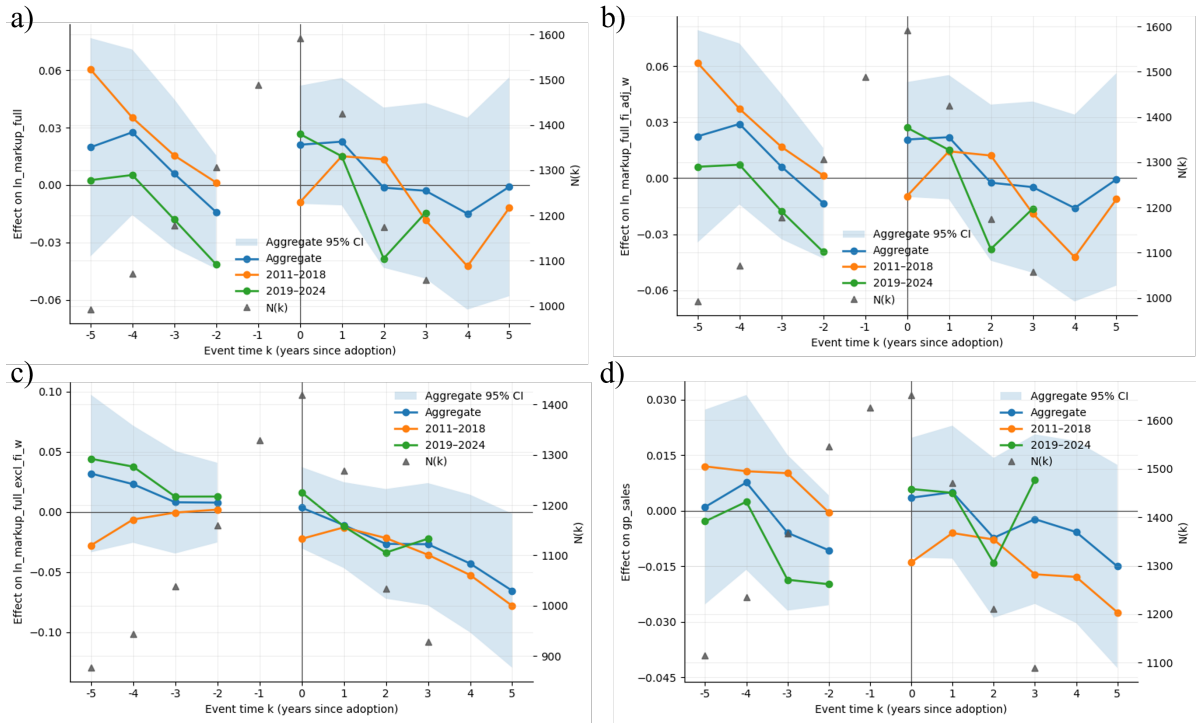


Figure 7: Pricing-power diagnostics around AI adoption. Event-time effects for log production-based markups under (a) a full-sample COGS-share construction, (b) a Finance & Insurance adjustment that adds interest expense to COGS for NAICS 52 firms when interest expense is observed before forming the variable-input revenue share, and (c) an exclusion of Finance & Insurance (NAICS 52) firms. Panel (d) also reports gross margin (gross profit over sales). Aggregate Sun–Abraham IW estimates with 95% confidence bands are shown along with adoption-vintage overlays.

R&D intensity increases mainly for early adopters. R&D/Assets rises in the aggregate, but the vintage overlay shows that the increase is concentrated among early adopters. Late adopters' R&D intensity is comparatively flat over the common-support horizon, consistent with a more plug-and-play adoption mode relying less on in-house development.

The intangible stock ramps up with a lag. The stock K^{intan} rises sharply after adoption, with effects growing over event time. This lagged accumulation is central for interpreting valuation dynamics: it is consistent with the idea that adoption creates option value and installed intangible capital that materializes over several years.

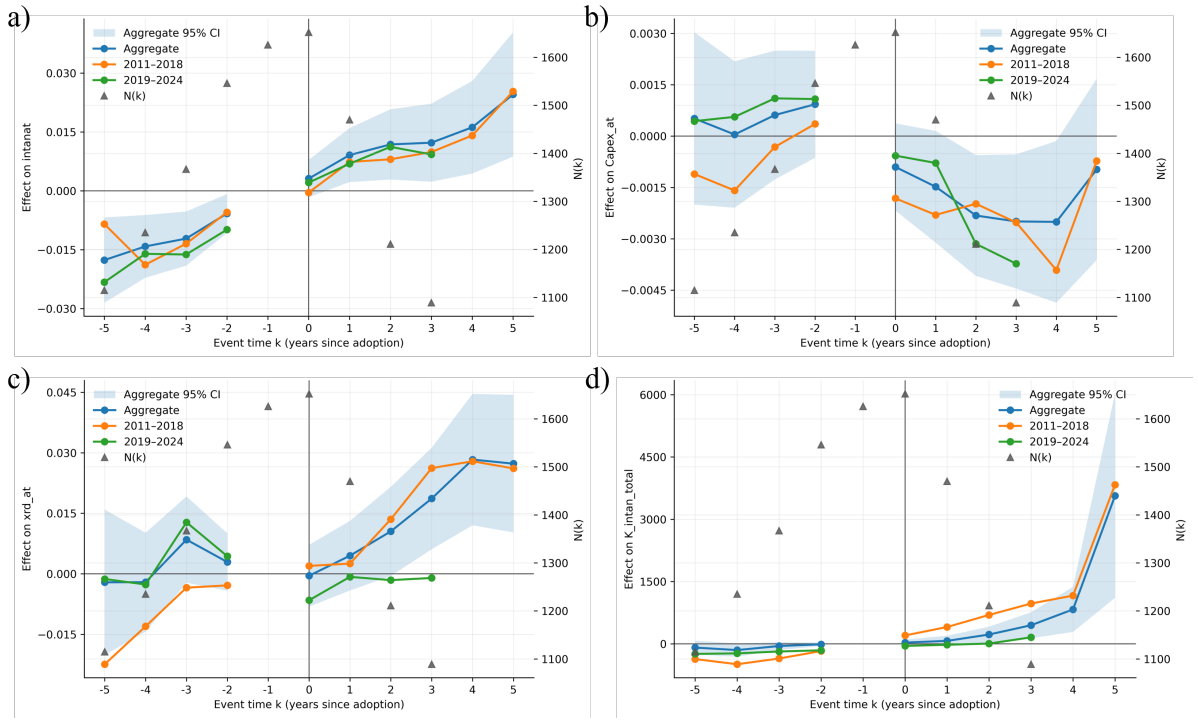


Figure 8: Investment and intangible deepening around adoption: Intangibles/Assets, Capex/Assets, R&D/Assets, and the stock of intangible capital K^{intan}.

4.2.5 Factor shares: reallocation away from labor

Figure 9 reports factor-share dynamics.[5] In value added, the capital share α_K^{VA} rises steadily after adoption while the labor share α_L^{VA} falls nearly one-for-one. In the GO3 decomposition, the intermediate-inputs (materials) share α_M^{GO3} drifts down modestly. An alternative BLS-style labor-share measure also declines in the medium run (though with wider confidence bands).

The key takeaway is within-firm reallocation: AI adoption is associated with a more capital- and intangible-intensive production structure and a lower labor share even though markups do not rise. This pattern is consistent with the interpretation that organizational and knowledge capital formation shifts the incidence of value added away from labor.

4.2.6 Valuation: a diffusion-dependent revaluation

We measure valuation using an asset-scaled Tobin's-q proxy and an enterprise-value-based q variant; Appendix C provides the exact constructions and data handling. Figure 10

[5]Appendix D.3 details construction of the wage-bill proxy, intermediate inputs and value added, and the α^{VA} and GO3 share mappings.

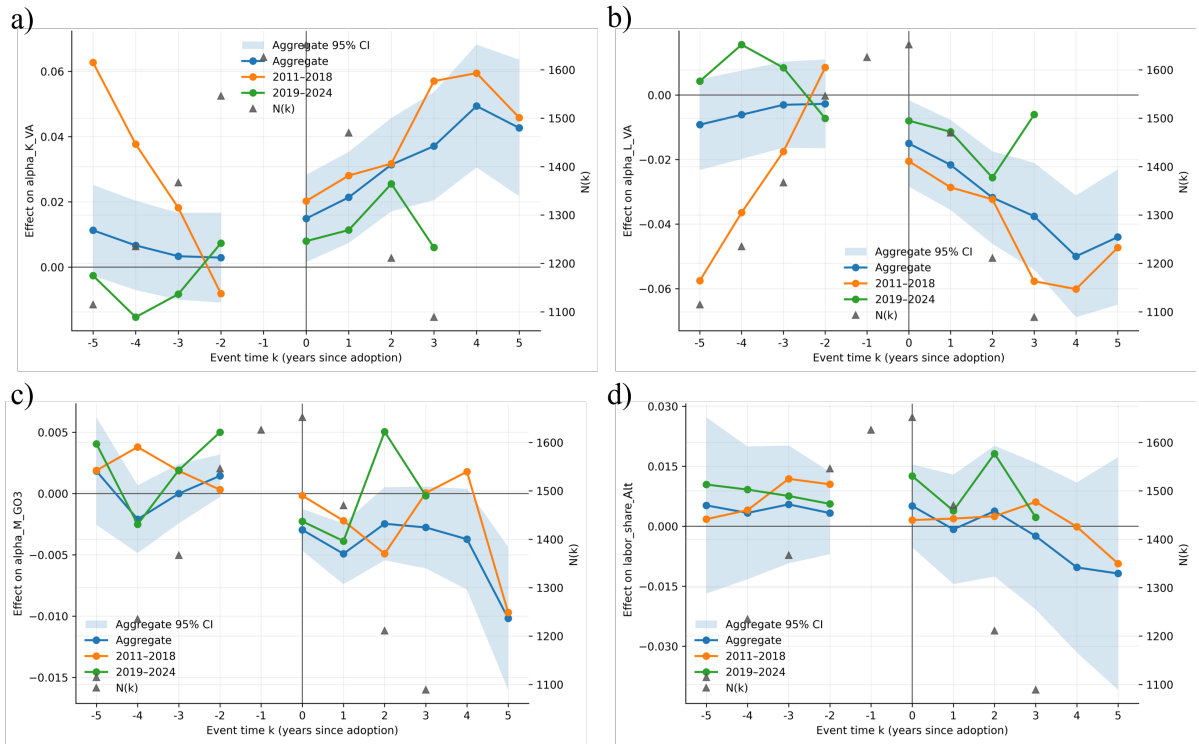


Figure 9: Factor-share dynamics around adoption: value-added capital and labor shares (α_K^{VA}, α_L^{VA}), GO3 materials share, and an alternative labor-share measure.

reports valuation dynamics using a Tobin's-q measure and an EV-based q. The aggregate estimates display a J-curve: valuations dip modestly around adoption and then rise strongly in the medium run.

The vintage overlay reveals the central result of the paper. Early adopters experience strong and increasing valuation premia after adoption, while late adopters experience weak or negative valuation responses over the same horizons. In other words, the market appears to price early adoption as a growth option and strategic capability, but it prices late adoption more as defensive catch-up with limited incremental rents.

4.3 Diffusion at adoption: first movers vs. followers

To isolate the role of diffusion at the time of adoption, we split adopters by the predetermined diffusion state in their industry. For each adopter i, define $S_i = S^{-i}_{g(i),G_i-1}$, the leave-one-out industry adoption share in the year prior to adoption. We classify i as a first mover if S_i lies below the 25th percentile among adopters within its industry; otherwise i is a follower. We then estimate interaction-weighted (IW) event studies separately for

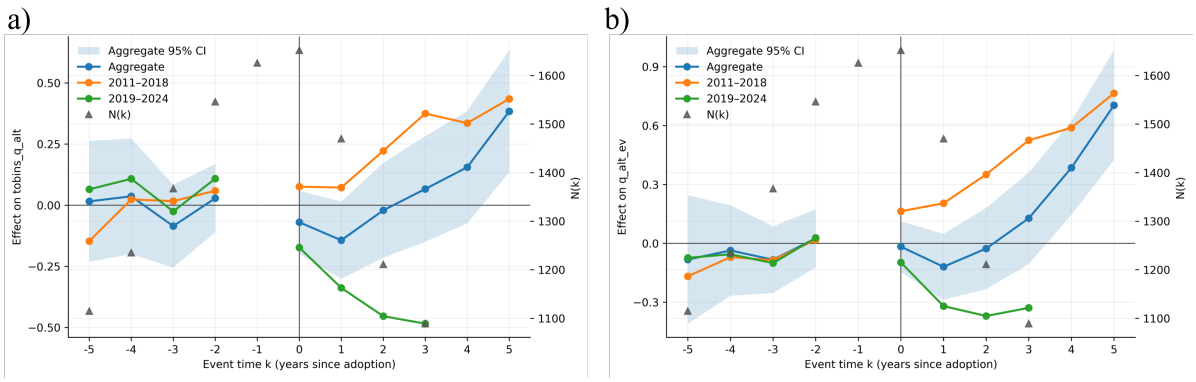


Figure 10: Valuation dynamics around adoption: Tobin's q_{alt} and EV-based Tobin's q (q_{alt}^{EV}). Aggregate IW estimates with 95% confidence bands, overlaid with adoption-vintage paths (early 2011–2018 vs. late 2019–2024).

first movers and followers (each compared to the same never-adopter pool), holding the support fixed at $k \in [-5, +5]$.

Figure 11 reports valuation dynamics for the two groups. Both groups exhibit a J-curve, but the timing and magnitude depend sharply on diffusion. Firms that adopt into low-diffusion states (first movers) revalue earlier and more strongly, with a smoother post-adoption rise. In contrast, firms that adopt after diffusion is already high (followers) display a deeper initial valuation dip and a smaller or later recovery within the event window. This pattern is consistent with a crowding interpretation: as AI becomes more common within an industry, adoption is closer to a parity technology and incremental rents are harder to sustain.

Appendix I shows that the first-mover/follower valuation split is robust to measuring diffusion two years prior to adoption ($S^{-i}_{g(i),\,G_i-2}$) and is directionally similar under a contemporaneous split ($S^{-i}_{g(i),\,G_i}$), though noisier.

Operating profitability responses do not mirror the valuation split. Figure 5 panels c and d show that first movers experience post-adoption operating-margin compression (point estimates drift downward for $k > 0$), consistent with larger adjustment costs (reorganization, learning, complementary investment) and/or stronger competitive responses that erode short-run margins. Followers exhibit a flatter margin path, with at most a modest and delayed improvement around mid-horizons. Given the much smaller support for first movers (left panel $N(k)$), these profitability patterns should be interpreted cautiously, but

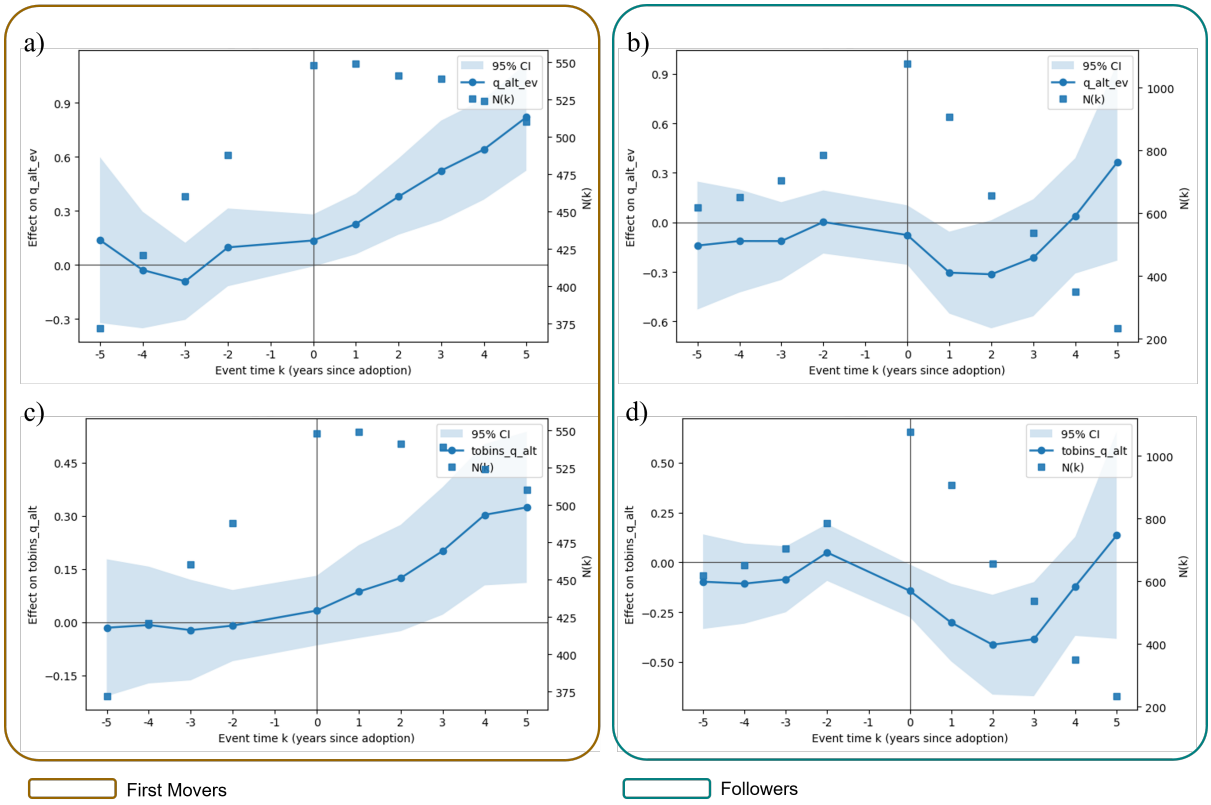


Figure 11: **Diffusion at adoption and valuation: first movers vs. followers (predetermined split).** Panels report Sun–Abraham interaction-weighted (IW) event-study estimates for valuation outcomes. Left panels plot first movers (adopters in low-diffusion states), and right panels plot followers (adopters in higher-diffusion states), each relative to the same never-adopter pool. First movers vs. followers are defined within industry using the bottom quartile of the predetermined diffusion state $S^{-i}_{g(i),G_i-1}$. Shaded regions are 95% confidence intervals; squares denote event-time support $N(k)$.

the joint evidence—earlier/stronger revaluation for first movers alongside weaker short-run margins—is consistent with markets pricing longer-run gains that are not immediately realized in accounting profitability.

Capability build-out and rent compression. Two empirical regularities sharpen the frontier-versus-catch-up interpretation. First, adoption vintages differ sharply in complementary capability accumulation: early adopters (2011–2018) exhibit a pronounced post-adoption rise in R&D intensity and in broad proxies for intangible/organizational-capital build-out, while later adopters (2019–2024) show a much weaker build-out profile within the common-support window. Book intangibles move far less across vintages, consistent with standard concerns that accounting intangibles under-measure internally created organizational capital. Second, valuation responses line up with diffusion-based

rent compression: early adopters display positive medium-run revaluation, while late adopters exhibit weaker or negative effects in the first few post-adoption years, consistent with incremental rents shrinking as industry diffusion rises. Appendix L reports additional evidence using predetermined splits (first movers vs. followers) and a diffusion×rents cross-sectional check that directly quantifies this rent-compression pattern.

4.4 Announcement vs. deployment

A central concern is reverse causality: firms may begin discussing AI or running limited pilots because performance is already improving. To probe this, we separate "pilot" activity from subsequent deployment using the discrete AI Adoption Score.

For each firm i, let $G_i^{0.5}$ denote the first year in which the AI Adoption Score reaches 0.5 (i.e., the first sustained pilot/announcement threshold). We then split these pilot adopters into two groups using end-of-sample adoption status: pilot-only firms that never transition to full deployment within our sample (score never reaches 1), and eventual deployers that later reach 1. We re-estimate Sun–Abraham event studies around $G_i^{0.5}$ for each group, comparing both to the same never-adopter pool and holding the event-time support fixed at $k \in [-5, +5]$.

Figure 37 shows a sharp contrast. Pilot-only firms exhibit muted and statistically imprecise responses in valuation and real outcomes throughout the event window. In contrast, eventual deployers display larger and more persistent post-event dynamics. This divergence is consistent with an implementation interpretation: the effects documented in our baseline results are tied to scaling and deployment rather than to AI rhetoric or exploratory activity at the pilot stage.

4.5 Level effects: vintage-heterogeneous 2SLS

The event studies document dynamic adjustment paths. We complement them with a panel IV design that estimates per-year level effects of being in the adopted state. We instrument the firm's strict AI adoption score with the lagged leave-one-out industry-year diffusion measure (Section 3). The pooled first stage is strong (robust F-statistic ≈ 44.9),

so diffusion shifts adoption meaningfully.

Because the event-study evidence shows sharp timing heterogeneity, we emphasize a vintage-heterogeneous 2SLS specification that allows effects to differ for early adopters ($G_i \leq 2018$) and late adopters ($G_i \geq 2019$). Table 3 reports the key outcomes most directly tied to our main three findings.

Table 3: Vintage-heterogeneous 2SLS effects of AI adoption on firm outcomes (early vs. late AI)

	Early AI (2011–2018)	**Late AI (2019–2024)**	N
Scale and cost structure			
Sales (USD millions)	11863.3**	2577.9*	19017
	(4639.6)	(1437.5)	
COGS / Sales	0.198***	0.096*	19017
	(0.058)	(0.050)	
SG&A / Sales	0.121**	-0.032	16626
	(0.049)	(0.055)	
Operating margin	-0.094***	-0.018	19017
	(0.035)	(0.032)	
Markup	-0.084	-0.110	18133
	(0.082)	(0.069)	
Markup (excl. financials)	-0.044	-0.101	17861
	(0.076)	(0.065)	
Markup (FI-adjusted)	-0.086	-0.110	18133
	(0.081)	(0.068)	
ln(Markup)	-0.210***	-0.157***	18133
	(0.069)	(0.059)	
ln(Markup) (excl. financials)	-0.195***	-0.157***	17861
	(0.070)	(0.059)	
ln(Markup) (FI-adjusted)	-0.209***	-0.158***	18133
	(0.069)	(0.059)	
Intangibles and investment mix			
Intangibles / Assets	0.055***	0.058***	18829
	(0.019)	(0.017)	
R&D / Assets	0.027**	0.005	18876
	(0.012)	(0.013)	
Capex / Assets	-0.003	-0.002	18989
	(0.003)	(0.003)	
Intangible capital K^{intan} (USD millions)	7618.1**	624.0**	16941
	(3030.5)	(263.6)	
Factor shares			
Capital share α_K^{VA}	0.093***	0.060***	19012
	(0.016)	(0.012)	
Labor share α_L^{VA}	-0.093***	-0.060***	19012
	(0.016)	(0.012)	
Capital share α_K^{GO3}	0.065***	0.037***	19012
	(0.011)	(0.008)	
Labor share α_L^{GO3}	-0.062***	-0.035***	19012
	(0.012)	(0.008)	
Valuation			
Tobin's q_{alt}	0.539**	-0.428*	19013
	(0.246)	(0.251)	
EV-based q_{alt}^{EV}	1.425***	-0.120	19017
	(0.343)	(0.424)	
Financing			
Debt / Assets (book leverage)	0.064**	0.018	18928
	(0.025)	(0.023)	
Cash / Assets	-0.070***	-0.040***	19017
	(0.018)	(0.014)	

Notes: Each row is a separate 2SLS regression with firm and year fixed effects and lagged controls. Endogenous regressors are cohort-specific AI adoption indicators (early vs. late), instrumented with the corresponding interactions of the lagged leave-one-out industry diffusion share. Standard errors are clustered by firm (reported in parentheses). Significance: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Notes:

Finding 1 in levels: cost shares rise; pricing-power proxies do not increase. The 2SLS estimates reinforce the event-study mechanism that adoption primarily operates through changes in scale and cost structure rather than through rising pricing power. For early compliers, Sales increases substantially and both COGS/Sales and SG&A/Sales rise, while operating margins decline by roughly 0.09 in ratio units. For late compliers, Sales also increases and COGS/Sales rises, but SG&A/Sales is statistically close to zero and operating margins are correspondingly less affected.

Across vintages, we find no evidence that pricing power increases with adoption. Markup levels are negative but imprecisely estimated, while ln(Markup) declines sharply and significantly for both early and late compliers, with similar magnitudes across the full-sample, Finance & Insurance adjusted, and Finance & Insurance exclusion constructions. Mechanically, because production-based markups are proportional to an estimated output elasticity divided by the flexible-input revenue share, the sizable post-adoption increase in COGS/Sales implies downward pressure on μ_{it} unless the inferred elasticity rises sufficiently to offset it. Taken together with the event-time markup diagnostics in Figure 7, the IV evidence supports an interpretation in which medium-run profitability and valuation dynamics are driven mainly by cost shares and investment composition (including intangible accumulation), rather than by an increase in markups.

Finding 2 in levels: intangible deepening and labor-share decline. Both early and late compliers exhibit large and precisely estimated increases in Intangibles/Assets and a corresponding shift in factor shares: α_K rises and α_L falls (in both VA and GO3 constructions). This is the level analogue of Figure 9 and aligns closely with the model's prediction that K^I/L rises after adoption and lowers the labor share even without higher markups.

Finding 3 in levels: valuation premia are concentrated among early adopters. The vintage IV results show a sharp valuation split. Early compliers experience a positive and significant increase in both Tobin's q and EV-based q. Late compliers exhibit a negative Tobin's-q effect (marginally significant) and a near-zero EV-q effect. This is the

cross-sectional counterpart of the diffusion-dependent valuation paths in Figures 10 and 11.

Robustness. A mechanical concern with diffusion-share designs is that diffusion may proxy for time-varying industry opportunity or industry-cycle timing, or be intertwined with fine-industry outcomes. We address this concern with three complementary checks. First, we replace the lagged diffusion shifter with the contemporaneous LOO diffusion share as a sensitivity check on timing (Appendix Table 16). Second, we augment the baseline specification with lagged industry-year aggregates (industry mean sales growth, profitability, and R&D intensity) to absorb predictable industry conditions (Appendix Table 17). Third, we reconstruct diffusion at a coarser industry aggregation to reduce simultaneity at very granular industry definitions (Appendix Table 18). Across these variations, the key qualitative patterns remain: early adopters are substantially larger and exhibit systematic shifts in cost structure and financing, while the most robust channel is the reallocation toward intangibles.

4.6 Heterogeneity by adoption mode: role and sourcing

We study whether dynamic effects of AI adoption differ by adoption mode. We focus on (i) whether the firm is primarily an AI user versus not, (ii) whether it is an AI vendor versus not, and (iii) whether adoption is primarily externally sourced versus not. To avoid conflating "not in group" with "unclassified/missing," we implement these splits on a classified-only sample: for each dimension we restrict to adopters for which the corresponding classifier is available, and then define the indicator $H_i \in \{0, 1\}$ within that classified subset.

Figure 12 summarizes the main heterogeneity patterns. Panels (a) and (b) show that post-adoption sales rise for all adopters, but effects are substantially larger for firms whose adoption is externally sourced and for firms classified as AI users. Panel (c) shows that externally sourced adopters also experience earlier and larger increases in COGS/Sales, consistent with external AI inputs and vendor services appearing on the cost side even as

revenues expand. Panel (d) shows that vendor-type adopters exhibit stronger long-run valuation gains, with the vendor group rising sharply at longer horizons. As usual in long-horizon event-time estimates, precision deteriorates in the tails as $N(k)$ falls, so k=4–5 estimates should be interpreted with appropriate caution.

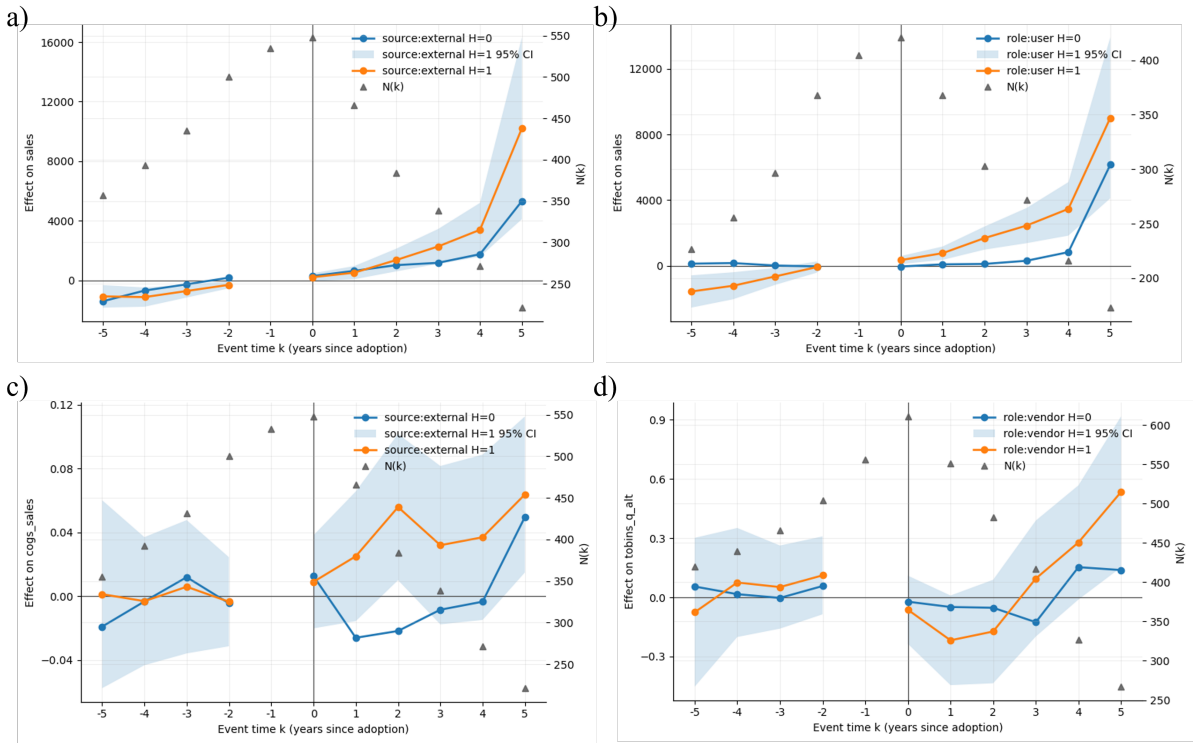


Figure 12: Heterogeneity by adoption mode (role and sourcing). **(a)** Sales by external sourcing. **(b)** Sales by user role. **(c)** COGS/Sales by external sourcing. **(d)** Alternative Tobin's q by vendor role. Samples are classified-only within each dimension. Triangles report $N(k)$.

Appendix J reports additional heterogeneity by vendor role for the production-function share objects.

4.7 Mechanisms: dose–response patterns by integration intensity

To probe mechanisms, we relate post-adoption dynamics to the time-varying intensity of AI integration. Unlike role/source indicators, integration intensity is continuous and evolves after adoption; thus we interpret these results as dose–response / channel evidence rather than ex-ante heterogeneity by predetermined firm type.

Let Z_{it} denote a standardized, time-varying integration measure (e.g., cumulative integration in logs). We report predicted event-time effects at $Z = 0$ (mean intensity)

and $Z = +1$ (one standard deviation above the mean), with uncertainty bands for the high-intensity path.

Figure 13 provides the key channel evidence. Panel (a) shows that higher integration intensity is associated with substantially larger declines in labor share. Panels (b) and (c) show that higher intensity coincides with stronger growth in intangible intensity (intangibles/assets) and in the level of total intangible capital. Taken together, the mechanism evidence suggests that deeper AI integration is accompanied by intangible capital deepening and a shift in factor shares away from labor, consistent with AI-enabled reorganization and automation channels.

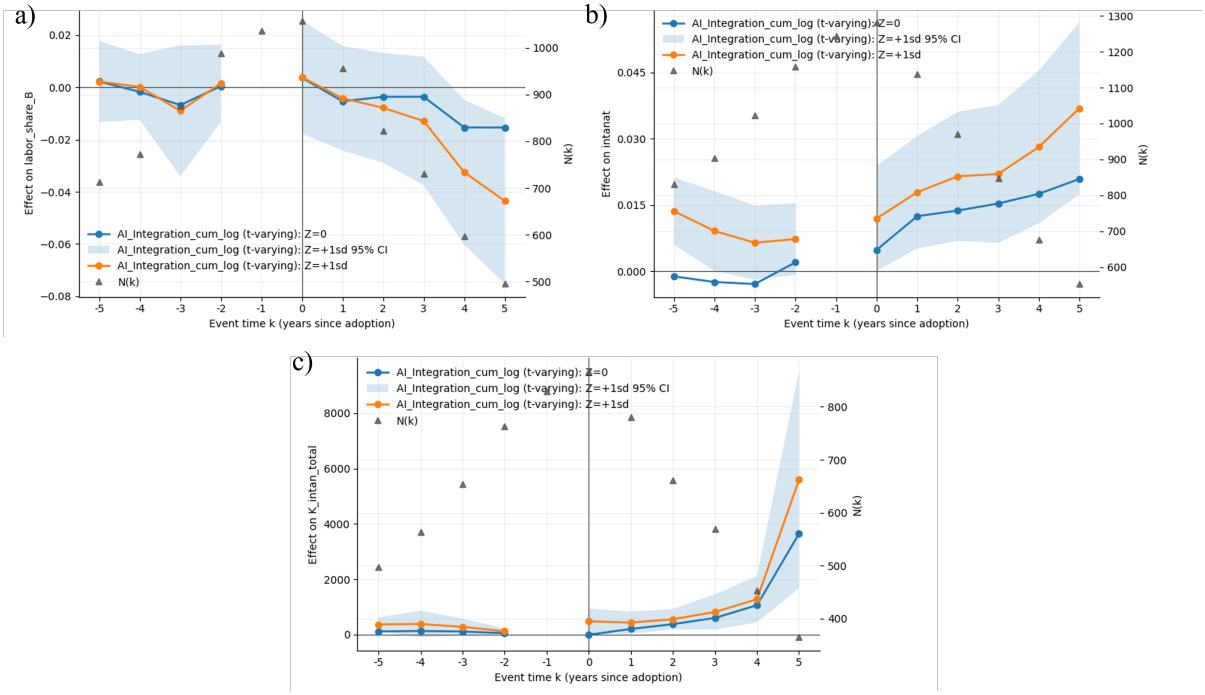


Figure 13: Mechanisms by time-varying AI integration intensity. Predicted event-time effects at $Z = 0$ and $Z = +1$ (one SD above mean intensity). **(a)** Labor share. **(b)** Intangibles/assets. **(c)** Total intangible capital. Triangles report $N(k)$.

Appendix K reports additional outcomes for the intensity-based mechanism exercise and shows a monotone valuation gradient with integration intensity under both intensity definitions.

5 Discussion and Interpretation

This section synthesizes the empirical patterns into a coherent interpretation and clarifies what the results do not imply. Throughout, the adoption year G_i should be interpreted as the first observable disclosure consistent with deployment/integration in investor communications, not necessarily the internal start date of experimentation.

5.1 AI as an intangible-intensive installation and retooling shock

Across specifications, the evidence aligns more closely with an "installation" narrative than with a pure market-power narrative. Around adoption, firms appear to undertake complementary intangible investments—data infrastructure, software and model development, and organizational redesign—that are largely expensed under standard accounting and therefore can depress (or fail to raise) conventional margins during the build-out phase. Over time, these investments coincide with scale expansion and intangible deepening. This interpretation is consistent with (i) rising intangible intensity and the growth of intangible-capital proxies after adoption, (ii) valuation and adjusted-profitability dynamics that strengthen in the medium run, and (iii) no robust sustained increase in production-based markup measures.

A key implication is measurement. Standard operating margins can understate the economic returns to AI-related investments during the build-out period when a substantial share of spending is expensed. Intangible-adjusted profitability metrics (Peters–Taylor style) are therefore informative for interpreting adoption dynamics, especially when comparing early versus late adopters.

5.2 Within-firm distribution: why labor shares fall even without higher markups

The factor-share results show that AI adoption is associated with a shift in value added away from labor and toward capital and intangibles. This pattern can arise even if markups do not rise. One channel is technological substitution and reorganization: AI can substitute

for routine tasks while complementing high-skill labor, reducing the aggregate labor share within the firm even if wages for some groups rise. A second channel is capital deepening in a broad sense: if AI raises the effective productivity of tangible and intangible capital, standard accounting decompositions mechanically allocate a larger share of value added to capital and intangibles. A third channel is measurement: organizational investments that appear as labor expenses in SG&A (e.g., engineers, consultants, managers) may be closer to capital formation economically; once these expenses are interpreted as investment (as in intangible-adjusted measures), the implied labor share can fall even without any increase in pricing power.

5.3 Why valuation gains are concentrated among early adopters

The sharp decline in valuation responses with diffusion is consistent with markets treating AI adoption as more value-relevant when it is rare and as closer to catch-up investment when it becomes common. When diffusion is low, early adopters may be better positioned to build proprietary data and workflow advantages, learn faster when organizational know-how is scarce, and capture temporary rents before competitive responses compress margins. As diffusion advances, adoption conveys less incremental information—investors may already expect broad adoption within the industry—and competitive catch-up can reduce the scope for sustained incremental rents. The predetermined diffusion splits (first movers versus followers) reinforce this pattern: adopters entering low-diffusion states revalue earlier and more strongly, while adopters entering high-diffusion states exhibit weaker or delayed revaluation within the event window.

5.4 Limitations and external validity

Three limitations are important when interpreting the results. First, the sample is U.S. public firms, which are larger and more intangible-intensive than the broader economy; external validity to private firms and small businesses is uncertain. Second, transcript-based adoption measures capture disclosed adoption; nondisclosure and strategic disclosure can generate measurement error and selection. Third, without worker-level data, the

results speak to labor shares and within-firm distribution at the accounting level rather than to wages, employment composition, or task-level impacts.

A related identification concern is that industry diffusion may proxy for industry-level opportunity or spillovers that directly affect outcomes. We mitigate this with lag structure, industry-level controls, and robustness to coarser diffusion definitions, but spillovers remain a natural limitation of diffusion-share designs. These limitations motivate future work linking transcript-based adoption to worker outcomes and extending measurement beyond public-firm disclosures.

6 Conclusion

This paper studies how the current wave of AI adoption reshapes firms' cost structure and within-firm factor shares, and how adoption is priced by equity markets. Using a new transcript-based measure of AI deployment timing and content for U.S. public firms (2011–2024), merged to Compustat outcomes, we document three robust patterns across interaction-weighted (IW) event studies and complementary fixed-effects IV specifications.

First, AI adoption appears primarily as an intangible-intensive installation and retooling shock rather than a sustained increase in pricing power. Production-based markups show no evidence of rising; operating cost shares and accounting margins move in ways consistent with expensed build-out and cost-structure reorganization. Second, adoption coincides with a reallocation of value added away from labor and toward capital and intangibles: intangible intensity rises, the constructed intangible stock accumulates, and within-firm labor shares decline. Third, the valuation payoff to adoption declines sharply with diffusion. Early adopters in low-diffusion environments experience substantially larger increases in Tobin's q than firms adopting after AI becomes widespread within their industry, consistent with an option-value interpretation of early adoption and with rent compression for late, catch-up adoption.

Overall, the evidence supports a view of AI as an intangible-heavy technology whose private returns depend critically on adoption timing and competitive diffusion, and whose

within-firm incidence operates through organizational capital and factor shares rather than markups alone. Future work can use the measurement approach developed here to link firm adoption to worker outcomes and task reallocation using matched employer–employee data, and to extend coverage beyond public firms.

References

Acemoglu, D. and Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of Labor Economics*, volume 4, pages 1043–1171. Elsevier.

Acemoglu, D. and Restrepo, P. (2018a). Artificial intelligence, automation, and work. In *The Economics of Artificial Intelligence*, pages 197–236. University of Chicago Press.

Acemoglu, D. and Restrepo, P. (2018b). The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review*, 108(6):1488–1542.

Acemoglu, D. and Restrepo, P. (2020). Robots and jobs: Evidence from US labor markets. *Journal of Political Economy*, 128(6):2188–2244.

Autor, D. H., Levy, F., and Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics*, 118(4):1279–1333.

Babina, T., Fedyk, A., He, A., and Hodson, J. (2023). Firm investments in artificial intelligence technologies and changes in workforce composition. *NBER Working Paper*, (31325).

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

Benkard, C. L., Miller, N. H., and Yurukoglu, A. (2025). The rise of market power and

the macroeconomic implications: Comment. Working Paper 34513, National Bureau of Economic Research.

Berlingieri, G., Calligaris, S., Criscuolo, C., and Verlhac, R. (2024). Laggard firms, technology diffusion, and its structural and policy determinants. *International Economic Review*.

Brynjolfsson, E. and McAfee, A. (2014). *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. W. W. Norton & Company, New York.

Brynjolfsson, E., Rock, D., and Syverson, C. (2017). Artificial intelligence and the modern productivity paradox. *NBER Working Paper*, (24001).

Corrado, C., Criscuolo, C., Haskel, J., Himbert, A., and Jona-Lasinio, C. (2021). New evidence on intangibles, diffusion and productivity. Working paper, OECD.

Corrado, C., Haskel, J., Jona-Lasinio, C., and Iommi, M. (2022). Intangible capital and modern economies. *Journal of Economic Perspectives*, 36(3):3–30.

De Loecker, J. and Warzynski, F. (2012). Markups and firm-level export status. *American Economic Review*, 102(6):2437–2471.

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

Eisfeldt, A. L., Schubert, G., and Zhang, M. B. (2023b). 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. *arXiv preprint arXiv:2303.10130*.

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

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.

Appendix

A Measurement, extraction, and validation

A.1 Transcript universe: counts and event-type composition

Table 4 reports annual transcript counts in the full corpus. Table 5 summarizes event types. The corpus is dominated by earnings calls, with substantial coverage of conference presentations and shareholder/analyst calls.

Table 4: Transcript counts by year

Year	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024
Events	8,928	10,167	9,992	10,249	10,507	10,561	12,846	13,232	13,675	16,783	18,218	17,430	17,866	17,455

Table 5: Event-type distribution in the full universe

Event type	Count
Earnings Calls	137,989
Company Conference Presentations	44,357
Shareholder/Analyst Calls	7,719
Special Calls	7,506
Analyst/Investor Day	3,698
M&A Calls	3,275
Sales/Trading Statement Calls	787
Guidance/Update Calls	520
Fixed Income Calls	101
Operating Results Calls	29
Interim Management Statement Calls	17
Sales/Trading Statement Release Date	4
Seeking Acquisitions/Investments	1

A.2 Keyword dictionary and normalization

We 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.

A.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.

A.3 Prompt schema and scoring examples

A.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.

A.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.

A.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.

A.3.4 Evidence-gated adoption scoring (extensive margin)

We 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 14 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.

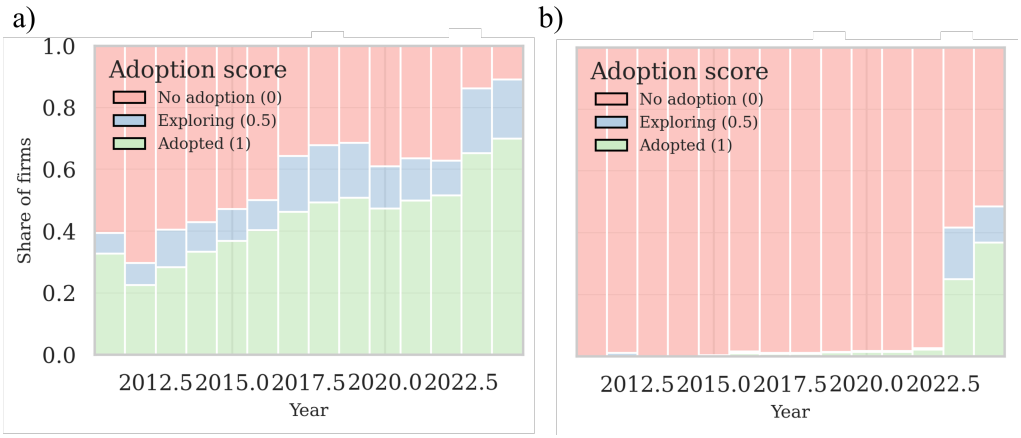


Figure 14: Adoption extent implied by evidence-gated scoring (AID), conditional on the LLM-coded sample. Stacked bars show, by year, the share of firm-years assigned $AID \in \{0, 0.5, 1\}$, where nonzero values require one or more transcript sentences as supporting evidence. The LLM is applied only to transcripts that contain AI-related keywords; therefore, the plotted shares are conditional on AI-mentioning firm-years and should not be interpreted as adoption rates in the full transcript universe. Panel A shows general AI adoption extent; Panel B shows GenAI adoption extent.

A.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.

A.3.6 Coverage distribution

Figure 15 shows the distribution of transcripts per firm–year in the merged analysis dataset.

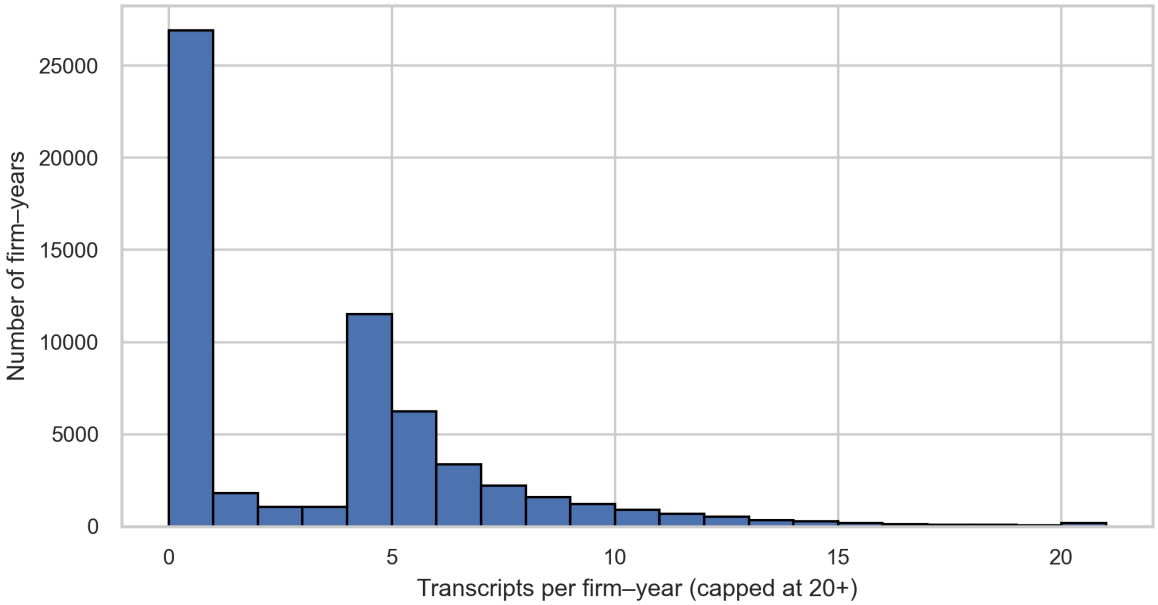


Figure 15: Distribution of transcripts per firm–year (capped at 20+). Most firm–years have 0–2 transcripts, with a long right tail.

A.3.7 Illustrative scoring cases

To make the evidence-gated scoring logic transparent, we 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, we 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 16 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 16: 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 17 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 17: 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).

A.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. We 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 6 summarizes the share of transcript–firm observations that include supporting excerpts by score class. As intended under evidence-gating, essentially all nonzero scores include supporting evidence, while the zero class contains no excerpt by design. Reporting these shares serves as a basic implementation check (i.e.,

that the pipeline enforced the evidence requirement as specified).

Table 6: Supporting evidence availability by score class

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

Automated anchoring of supporting excerpts (evidence locatability). A key failure mode in LLM pipelines is fabricated supporting text. Our prompts instruct the model to return an excerpt drawn from the transcript (allowing minor formatting differences), and we 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, we use a hierarchical matching procedure. First, we attempt normalized substring matches after case-folding and whitespace normalization. For remaining cases, we 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 7 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 7: 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, we 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 8 reports abstention rates in the keyword-screened, parsed transcript sample ($N = 23{,}459$). 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 8: 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. We 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). We treat the human label as the reference and compare it to the LLM score. Tables 9 and 10 report confusion matrices (LLM score in rows; human score in columns), along with class-wise precision, recall, and F1.

Table 9: 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 9). 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.

Table 10: 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).

For GenAI Adoption, agreement is higher: overall three-class accuracy is 0.967 (Table 10). 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.

B Additional descriptive evidence

B.1 Sourcing, strategy, applications, technologies, and business units

B.1.1 Implementation depth (Intensive margin)

Figure 18 summarizes the disclosure-based AI integration level used in the mechanism analysis. The measure is ordinal (Levels 1–3) and captures whether managers describe AI as (i) limited or early deployment, (ii) operationally deployed in at least one function or product line, or (iii) broadly integrated into core operations and/or products.



Figure 18: AI integration level over time (Intensive margin). Heatmap entries report the annual count of AI-screened events classified into Levels 1–3 based on disclosure evidence.

B.1.2 Sourcing

Table 11 and Figure 19 summarize sourcing among AI-screened events where disclosures are informative about whether AI capabilities are built in-house, sourced externally, or use a mixed approach. In-house development is the modal category in this sample. External and mixed sourcing become more prevalent in later years, consistent with the expansion of vendor ecosystems and third-party foundation-model offerings.

Table 11: AI sourcing in the screened sample. Entries report counts and shares for in-house, external, and mixed (both) sourcing among events with identifiable sourcing disclosures.

Source	Count	Share (%)
In–house	10,246	81.3%
External / Others	1,626	12.9%
Both	725	5.8%

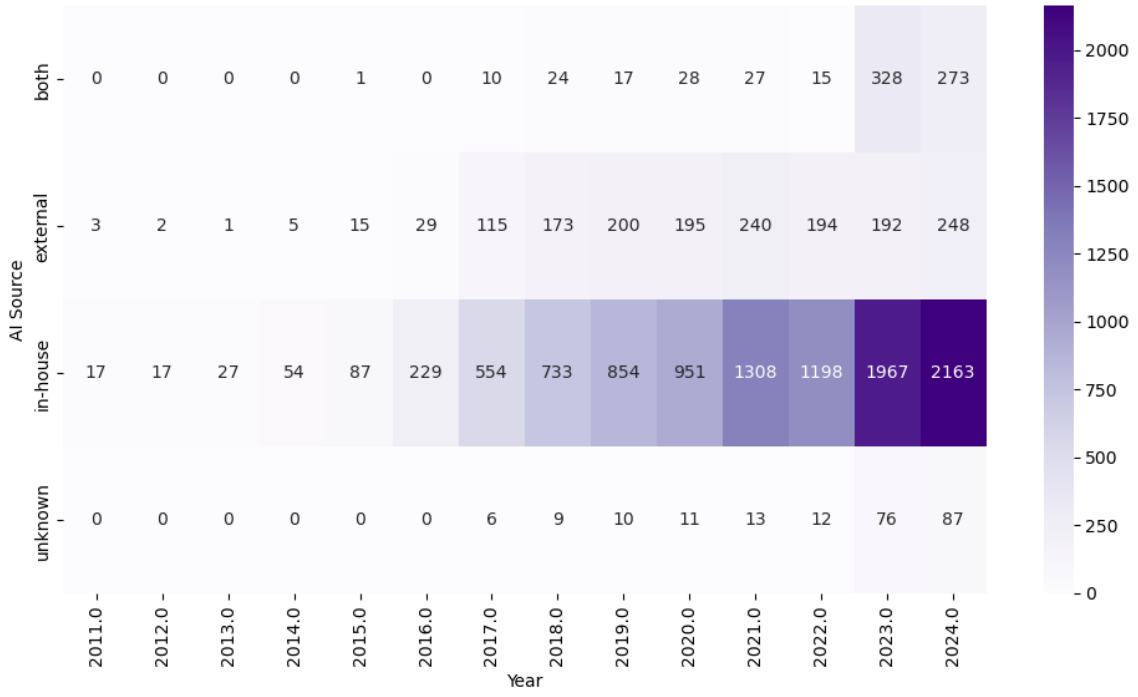


Figure 19: AI sourcing by year. Counts reflect annual mentions classified as in-house, external, both, or unknown based on disclosure text.

B.1.3 Strategic positioning

Figure 20 reports whether managers frame AI as a core strategic driver versus a supportive tool. Both framings rise over time, with a noticeable increase after 2020 and especially

in 2023–2024, indicating that AI increasingly appears as a recurring element of strategy discussion rather than sporadic experimentation.

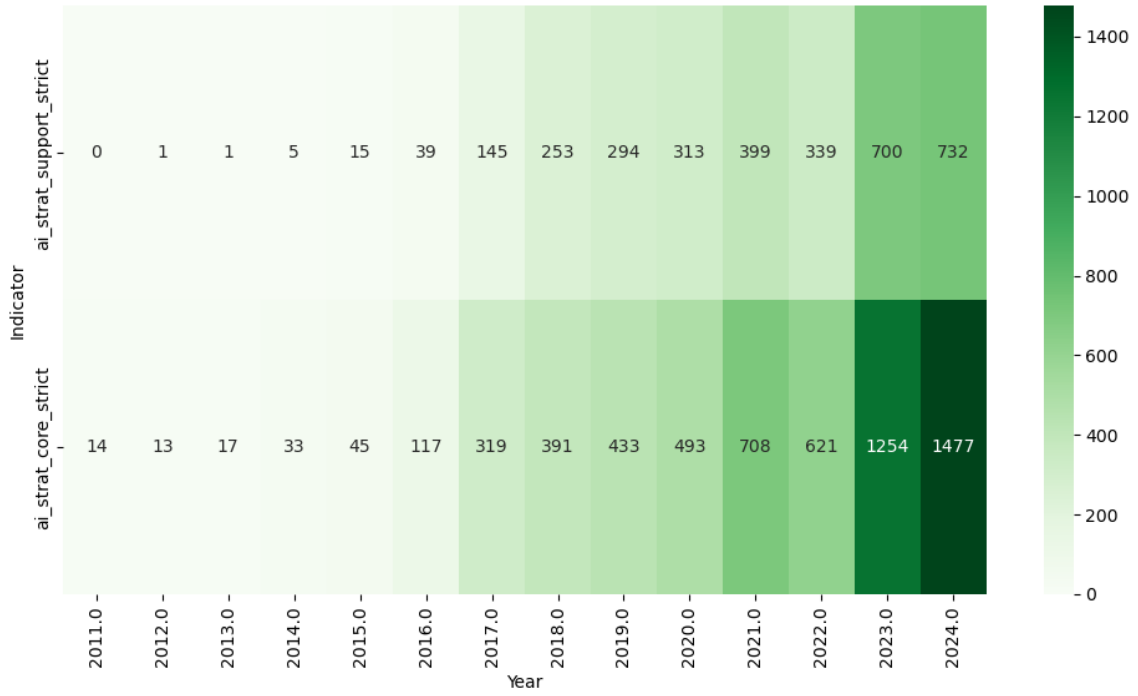


Figure 20: Strategic positioning by year. Heatmap entries report annual counts of AI-screened events in which managers frame AI as a core strategic driver versus a supportive tool.

B.1.4 Application types

Figure 21 decomposes AI-related events by inferred application type. Analytical/predictive applications are the most frequently identified category throughout, consistent with earlier diffusion of forecasting, optimization, and monitoring use cases. Generative/creative and interactive applications rise sharply in 2023–2024, aligned with LLM-enabled products and workflows. Autonomous/robotics use cases remain smaller in count, consistent with higher deployment frictions in physical settings.

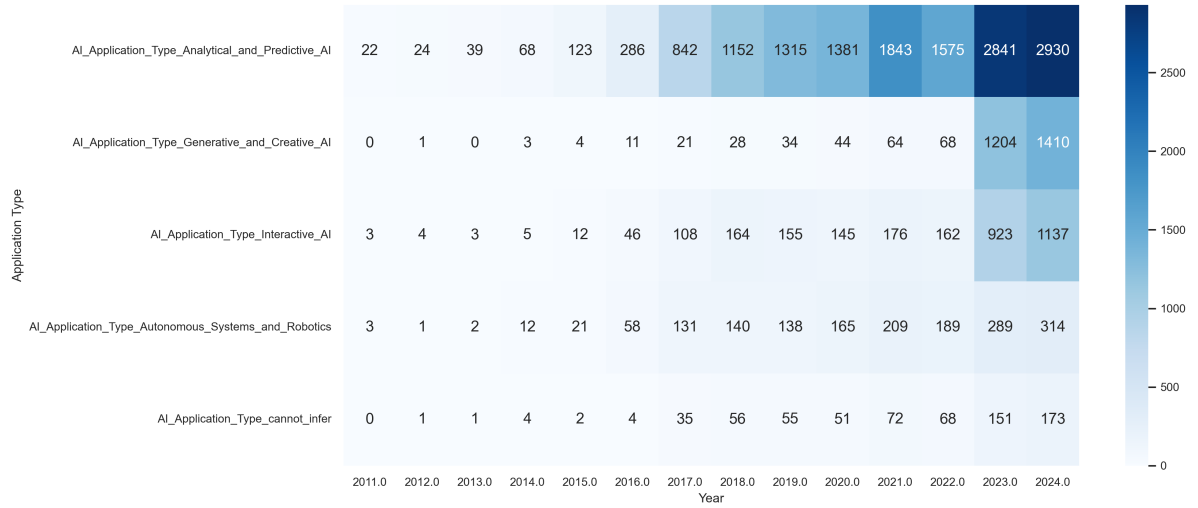


Figure 21: Application types by year. Heatmap entries report the annual count of AI-screened events classified by inferred application type from disclosure text.

B.1.5 Business-unit context

Figure 22 reports inferred business-unit context for AI deployment. Mentions increase across all major units over time, with the largest counts in product development/R&D and operations, followed by IT/cybersecurity and sales/marketing. The broadening across units suggests that disclosures increasingly describe operational and product-facing implementations rather than purely IT-scoped initiatives.

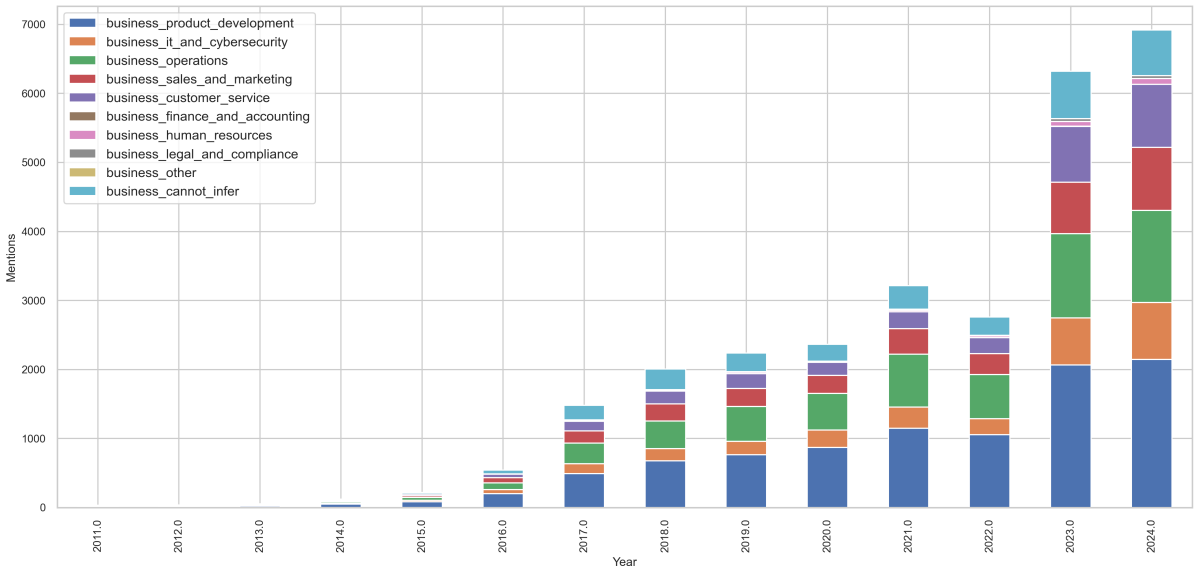


Figure 22: Business-unit references by year: product/R&D, IT/security, operations, sales/marketing, customer service, finance/accounting, HR, legal/compliance, other.

B.1.6 Sentiment of disclosed AI impact

Figure 23 summarizes the sentiment of disclosed AI impact among adoption events. Disclosures are overwhelmingly positive: approximately 91% of adoption events are coded as positive and the remainder as neutral, with essentially no negative sentiment in this sample. This skew is consistent with incentives in public communications. Accordingly, the measure should be interpreted as managerial framing conditional on discussing adoption, not as a welfare assessment or a summary of realized outcomes.

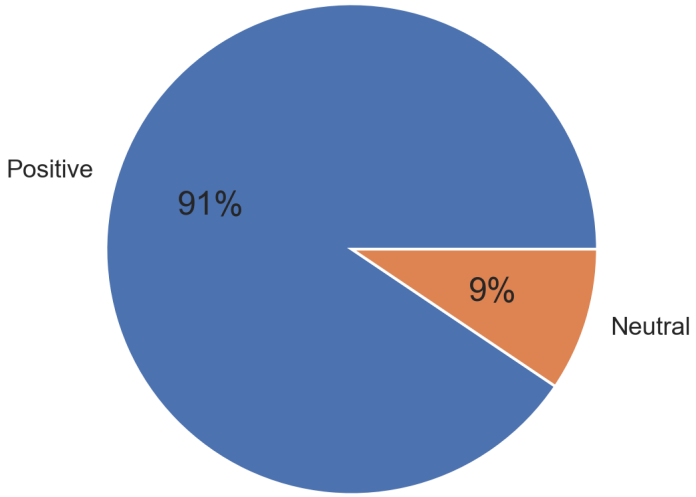


Figure 23: Overall sentiment of AI impact among adoption events.

B.1.7 Technology mentions over time

Figure 24 reports the time series of technology categories referenced in AI-related events. The rise after 2022 reflects both increased diffusion and the emergence of foundation-model language (e.g., GenAI/LLMs). Categories are not mutually exclusive; a single event can contribute to multiple counts.

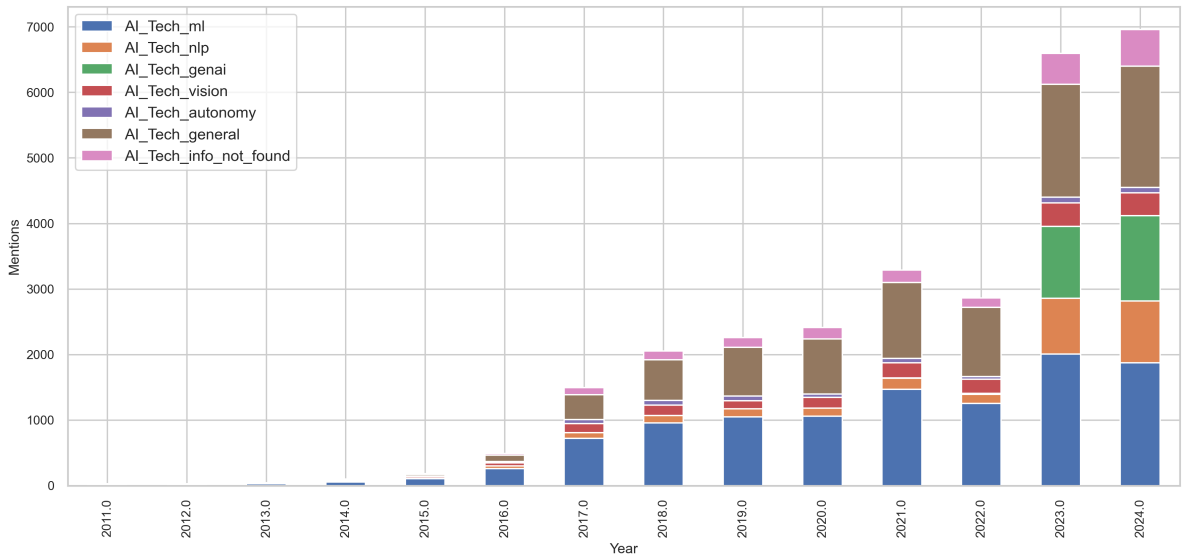


Figure 24: AI technology mentions by year. Stacked bars report annual counts of disclosure mentions mapped to technology categories (ML, NLP, GenAI, computer vision, autonomy, general AI, and no specific technology identified). Categories are not mutually exclusive.

B.2 Industry composition and diffusion

B.2.1 Industry composition of adopters

Figure 25 summarizes the industry composition of AI-adopting firms. Adoption events are concentrated in Software & Services, followed by Health Care Equipment & Services, Technology Hardware, Semiconductors, Financial Services, and Commercial & Professional Services. The concentration in software is expected given both supply-side AI development and earlier feasibility of deployment in digital production processes; however, adoption spans manufacturing, consumer, and regulated sectors, consistent with broad diffusion beyond pure technology industries.

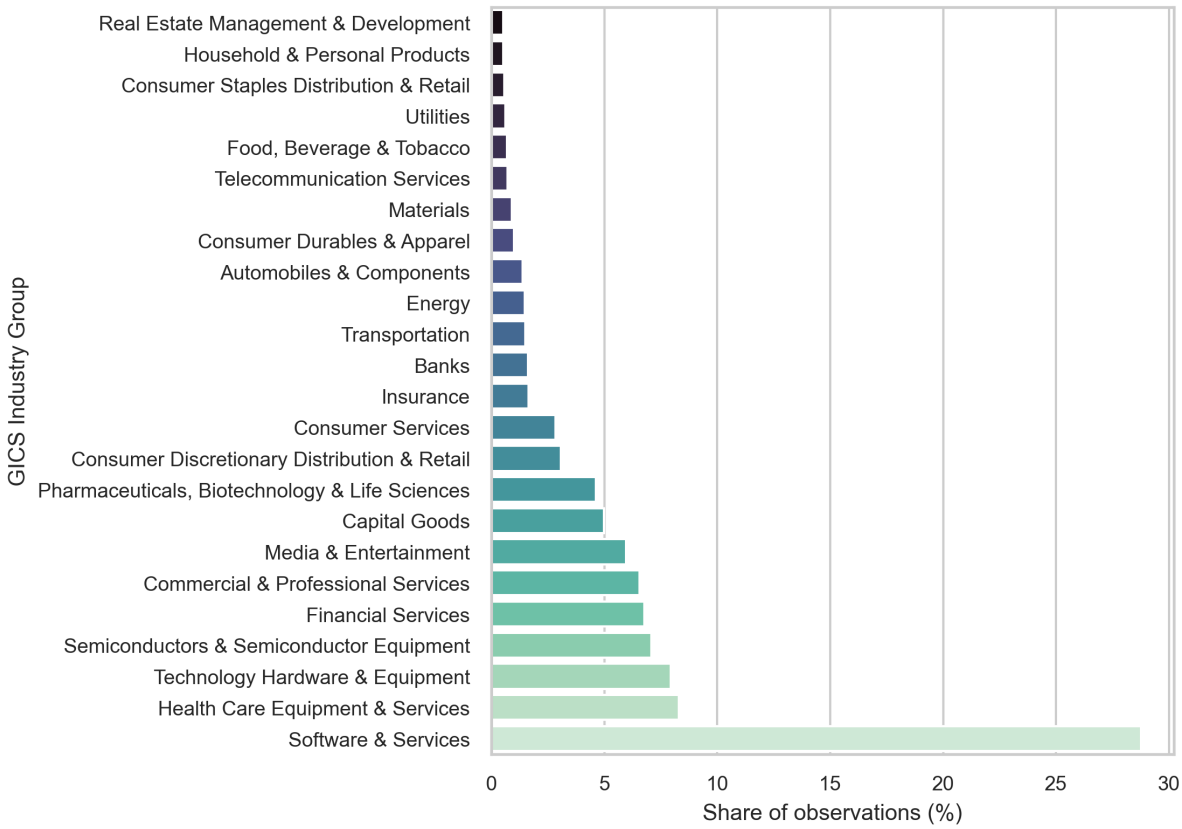


Figure 25: Distribution of AI-adopting firms by GICS Industry Group.

B.2.2 AI adoption heatmap

Figure 26 reports within-industry adoption rates over time. Adoption rises materially after 2017–2018 across most industry groups. Diffusion is heterogeneous: Software & Services reaches the highest adoption rates by 2023–2024, while utilities, energy, and basic materials exhibit slower take-up. This heterogeneity motivates our emphasis on timing and diffusion state in the main analysis.

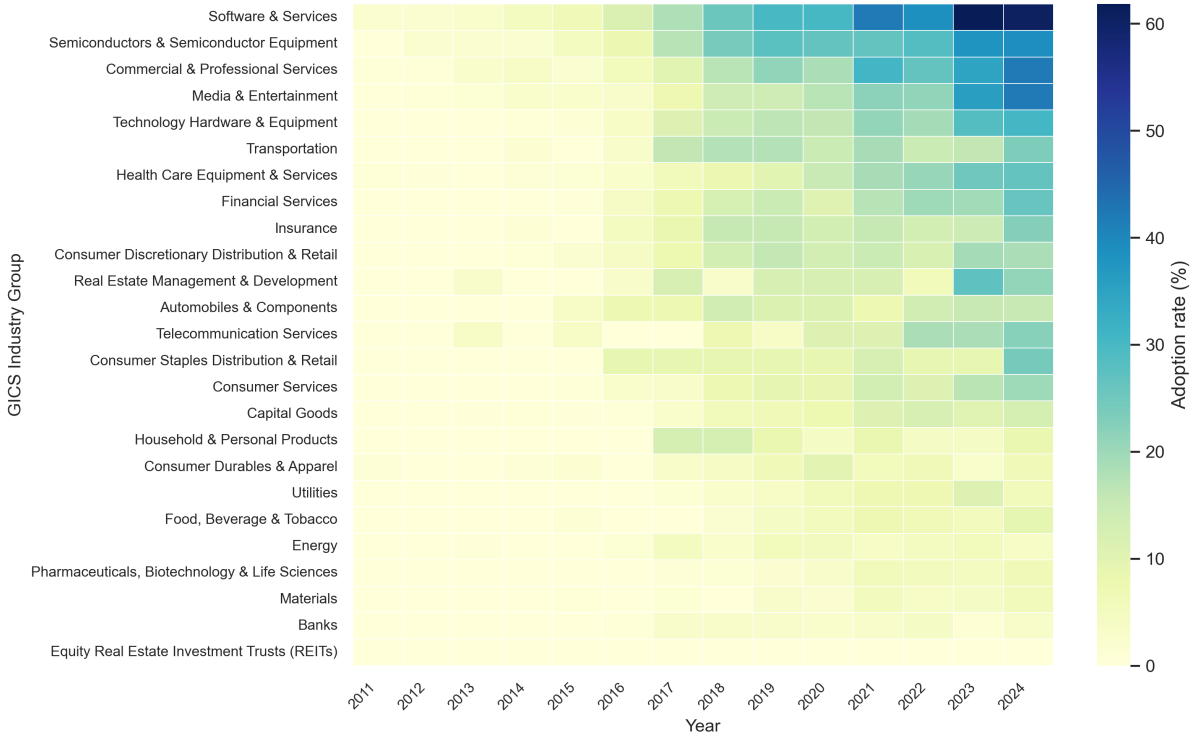


Figure 26: AI adoption rate (%) by GICS Industry Group and year.

B.2.3 GenAI adoption heatmap

Figure 27 reports within-industry GenAI adoption rates over time. Diffusion is back-loaded: adoption is near zero through 2021, becomes visible in 2022, and rises sharply in 2023–2024. Uptake is concentrated in a small set of information- and software-intensive sectors. Software & Services exhibits the highest GenAI adoption by 2023–2024, with additional increases in Media & Entertainment and Commercial & Professional Services. A second tier (e.g., Semiconductors, Financial Services, Technology Hardware) shows meaningful but lower uptake, while many remaining sectors remain close to zero even by 2024. Overall, the heatmap underscores that GenAI represents a distinct and more recent diffusion wave than earlier AI.

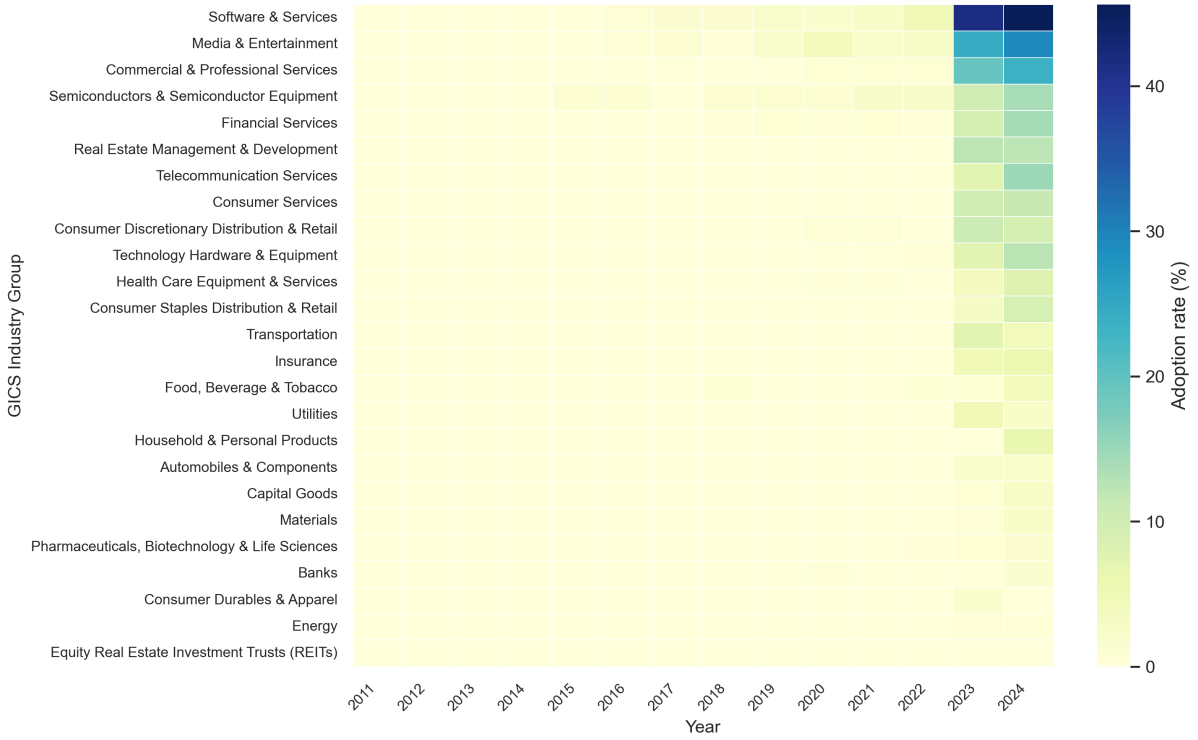


Figure 27: GenAI adoption rate (%) by GICS Industry Group and year.

B.2.4 Diffusion dynamics: adoption S-curves

Figures 28 and 29 plot adoption rates over time across GICS industries at different levels of aggregation. Diffusion is nonlinear: adoption remains low through the mid-2010s and accelerates later in the sample. The timing also differs across AI waves: broader AI adoption accelerates earlier, while GenAI diffusion is concentrated in 2023–2024. These dynamics motivate the early/late vintage split (2011–2018 vs. 2019–2024) and support using lagged leave-one-out industry adoption shares as instruments capturing peer-driven diffusion pressure.

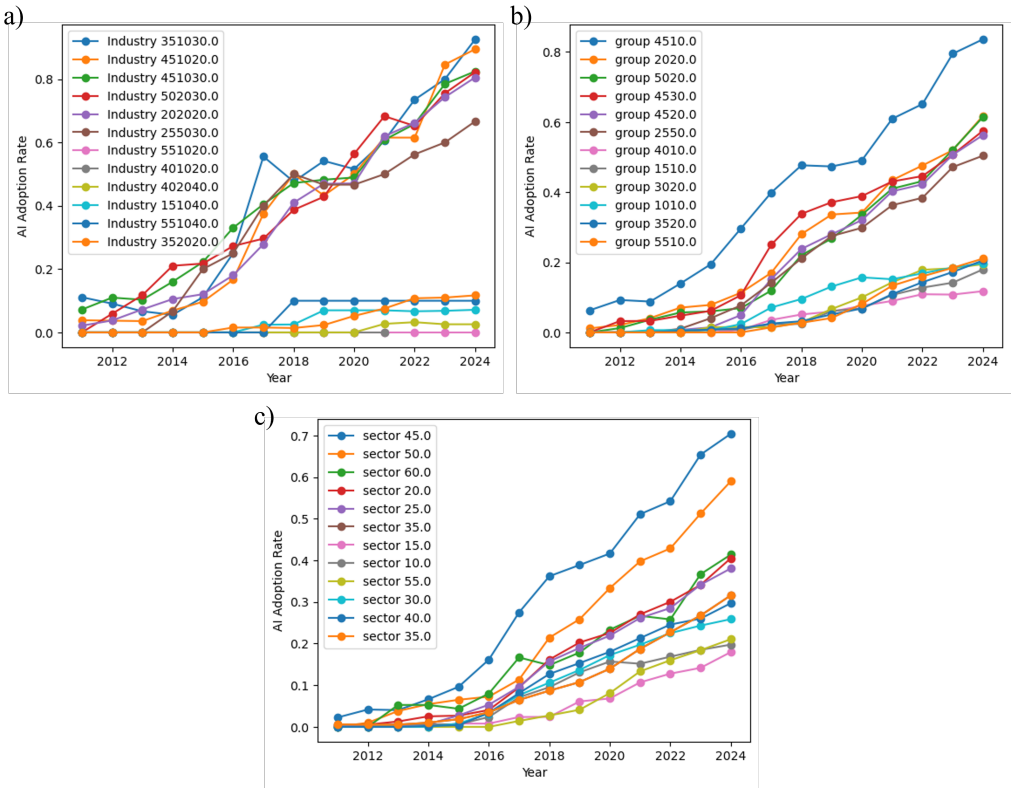


Figure 28: AI adoption diffusion over time across industries and sector aggregates.

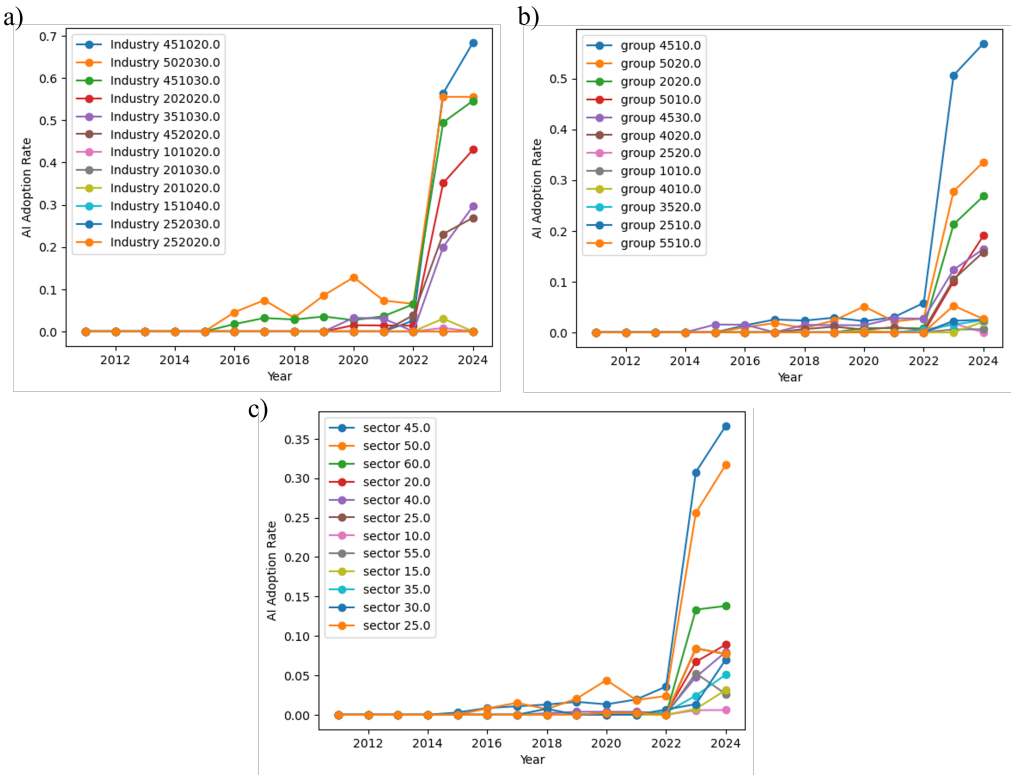


Figure 29: GenAI adoption diffusion over time.

B.3 External vendor landscape

A key advantage of transcript-based extraction is that it can recover not only whether firms adopt AI, but also how they source it. When managers explicitly mention external AI providers (e.g., model developers, cloud platforms, enterprise AI stacks), the structured fields record vendor names, allowing the construction of vendor-level adoption networks and vendor–industry flows. The figures below illustrate these objects and summarize the external AI ecosystem present in disclosures.

B.3.1 Vendor concentration, overlap, and industry flows

Figure 30 reports the top external vendors referenced as AI providers by number of distinct client companies. The distribution is top-heavy, consistent with a small set of general-purpose model and cloud providers serving a broad client base.

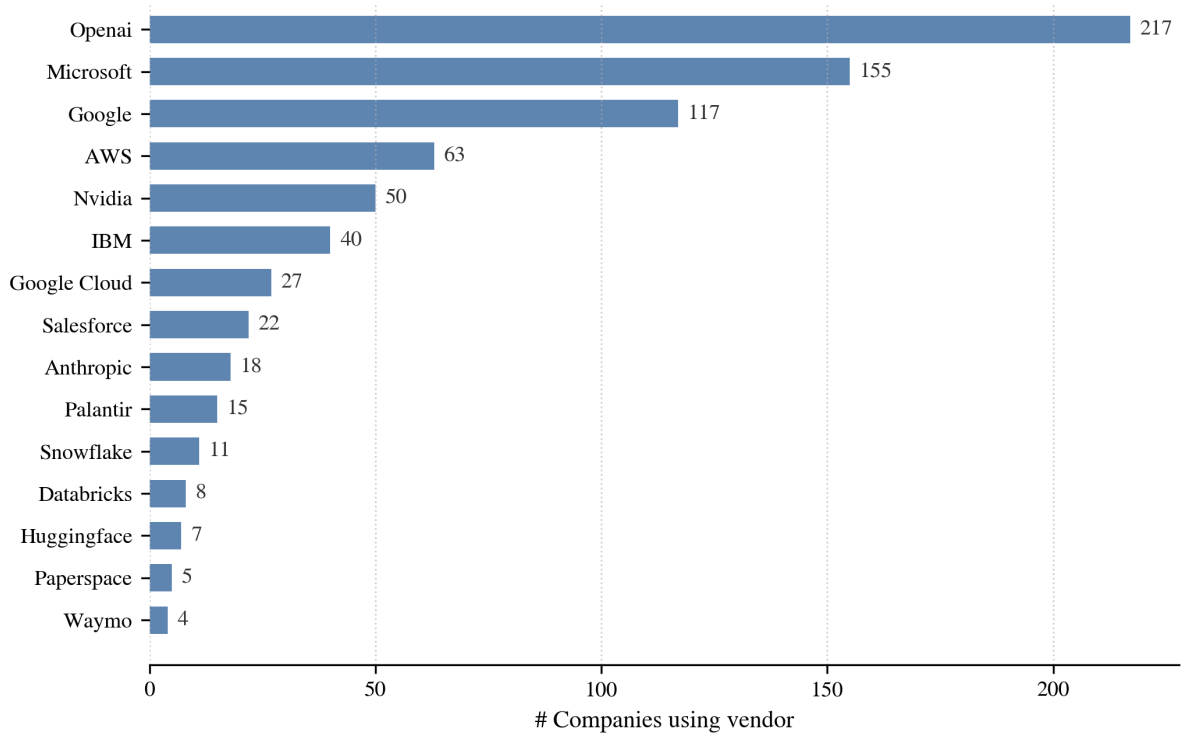


Figure 30: Top AI vendors by number of client companies.

Figure 31 visualizes vendor co-usage: nodes are vendors and edges indicate at least five shared clients. The overlap among leading vendors is consistent with multi-vendor AI stacks (e.g., combining cloud infrastructure, foundation models, and enterprise platforms).

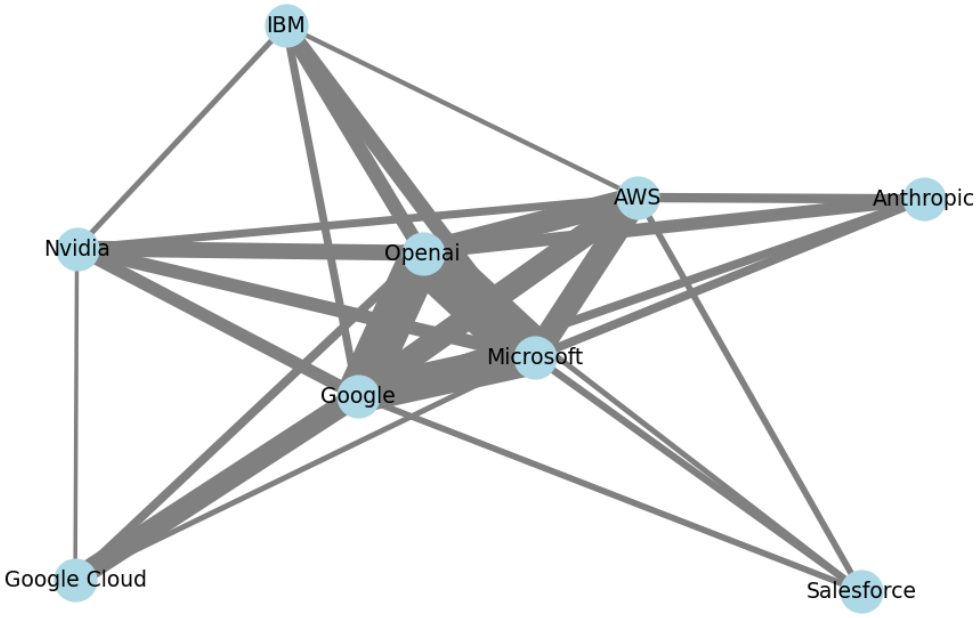


Figure 31: Vendor–vendor overlap network (threshold ≥ 5 shared clients).

Figure 32 summarizes vendor-to-industry flows for the top vendors. The largest destination is Software & Services, but flows extend across many industry groups, illustrating cross-sector external sourcing.

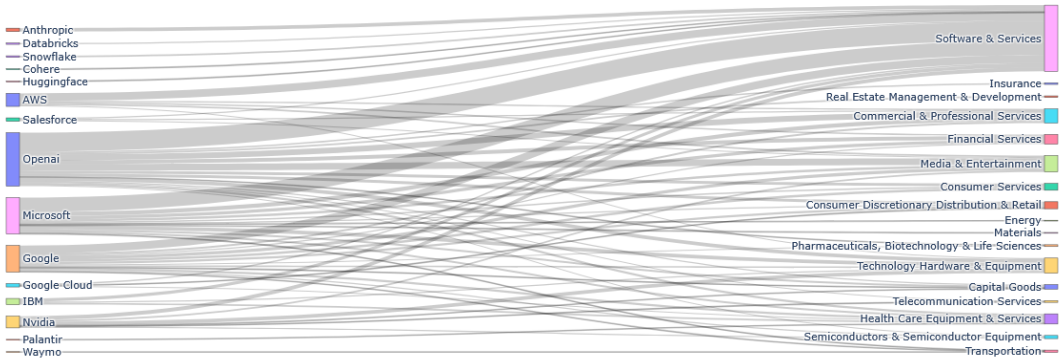


Figure 32: Vendor $\rightarrow$ 4-digit GICS group flows (Top 20 vendors, flows ≥ 3).

Figure 33 presents a company–company projection network where edges indicate shared vendor usage. This object can be interpreted as a "stack similarity" graph and may be useful for studying diffusion channels, peer effects, or clustering of AI stacks.

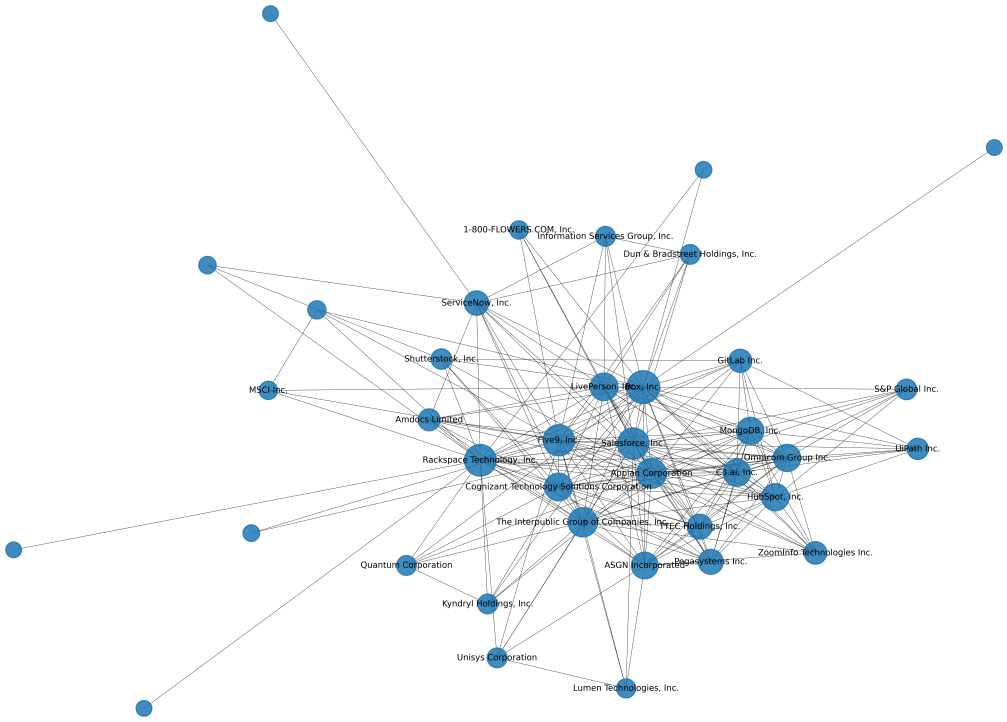


Figure 33: Company–company projection of shared AI vendors.

C Data sources and variable construction

Firm outcomes are from Compustat North America annual fundamentals and are indexed by gvkey and fiscal year t. Market equity is measured as fiscal-year price times shares outstanding (using Compustat when available). All variables are constructed at the firm–year level.

Core outcomes. We study operating cost shares and valuation. Operating margin is operating income after depreciation scaled by sales, $\text{OIADP}_{it}/\text{Sales}_{it}$. Cost shares are $\text{COGS}_{it}/\text{Sales}_{it}$ and $\text{XSGA}_{it}/\text{Sales}_{it}$. The book intangible ratio is intan/at when observed, and R&D and capex intensities are $\text{XRD}_{it}/\text{AT}_{it}$ and $\text{CAPX}_{it}/\text{AT}_{it}$. We use two asset-scaled valuation proxies:

$$\text{Tobin's q}_{it} = \frac{\text{AT}_{it} + MVE_{it} - \text{CEQ}_{it}}{\text{AT}_{it}}, \qquad \text{q}_{it}^{EV} = \frac{MVE_{it} + \text{PS}_{it} + \text{DLTT}_{it} + \text{DLC}_{it}}{\text{AT}_{it}},$$

setting preferred stock to zero when missing.

Treatment timing and diffusion. Let $\text{AID}_{i\tau} \in \{0, 0.5, 1\}$ denote the evidence-gated adoption score assigned to transcript τ of firm i (Appendix A.3). We aggregate transcript-level evidence to firm–year indicators as follows. The strict adoption indicator is

$$\text{AI_score_strict}_{it} \equiv \mathbb{1}\left\{\max_{\tau \in \mathcal{T}_{it}} \text{AID}_{i\tau} = 1\right\},$$

i.e., it equals one if firm i has at least one transcript in fiscal year t with confirmed adoption (AID = 1). The lenient indicator is

$$\text{AI_score_lenient}_{it} \equiv \mathbb{1}\left\{\max_{\tau \in \mathcal{T}_{it}} \text{AID}_{i\tau} \geq 0.5\right\},$$

i.e., it equals one if any transcript in year t reflects at least exploratory adoption language (AID ≥ 0.5). The lenient definition is used as a robustness check; results are qualitatively similar.

The adoption year G_i is defined as the first fiscal year in which the baseline indicator switches on, $G_i = \min\{t : \text{AI_score_strict}_{it} = 1\}$. Event time is $k = t - G_i$, with $k = -1$ as the omitted reference period (top- and bottom-coding as indicated in each figure). Industry diffusion is measured using the lagged leave-one-out share $S^{-i}_{g(i),t-1}$ defined in the main text/IV section.

Data handling. Ratios are formed after setting zero denominators to missing. Unless otherwise noted, variables are winsorized by fiscal year at the 1st and 99th percentiles.

D Outcome construction details

D.1 Intangible capital and intangible-adjusted operating profitability

We construct intangible-capital stocks and an intangible-adjusted operating profitability measure following Peters and Taylor (2017). Let XRD_{it} denote R&D expense and XSGA_{it} denote SG&A. We treat XRD_{it} as investment in knowledge capital and a fraction λ_{SGA} of XSGA_{it} as investment in organizational capital. The baseline sets $\lambda_{\text{SGA}} = 0.30$ and we vary it in robustness exercises.

Perpetual-inventory stocks. Define knowledge-capital and organizational-capital stocks, K^R_{it} and K^O_{it}, with depreciation rates δ_R and δ_O:

$$K^R_{it} = (1-\delta_R)K^R_{i,t-1} + \text{XRD}_{it}, \qquad K^O_{it} = (1-\delta_O)K^O_{i,t-1} + \lambda_{\text{SGA}}\text{XSGA}_{it}.$$

The first observed stock for each firm is initialized using a steady-state approximation,

$$K^R_{i,t_0} = \frac{\text{XRD}_{i,t_0}}{\delta_R}, \qquad K^O_{i,t_0} = \frac{\lambda_{\text{SGA}}\text{XSGA}_{i,t_0}}{\delta_O}.$$

We set $\delta_R = 0.20$ and $\delta_O = 0.33$ in the baseline. Total intangible capital is $K^{\text{intan}}_{it} = K^R_{it} + K^O_{it}$. When XRD_{it} or XSGA_{it} is missing, the corresponding investment flow is set to zero in the stock update.

Economic amortization and adjusted operating income. Economic amortization is

$$\text{Amort}^{\text{intan}}_{it} = \delta_R K^R_{it} + \delta_O K^O_{it}.$$

Because OIADP_{it} is measured after expensing R&D and SG&A, we construct intangible-adjusted operating income by capitalizing the investment components and subtracting

economic amortization:

$$\mathrm{OIADP}^{*}_{it} = \mathrm{OIADP}_{it} + \mathrm{XRD}_{it} + \lambda_{\mathrm{SGA}}\mathrm{XSGA}_{it} - \mathrm{Amort}^{\mathrm{intan}}_{it}.$$

The corresponding intangible-adjusted operating margin is

$$\mathrm{OperMargin}^{*}_{it} = \frac{\mathrm{OIADP}^{*}_{it}}{\mathrm{Sales}_{it}}.$$

D.2 Markups and price–cost margins

Gross margin / price–cost margin. We report a gross-margin proxy (price–cost margin),

$$\mathrm{PCM}_{it} = \frac{\mathrm{Sales}_{it} - \mathrm{COGS}_{it}}{\mathrm{Sales}_{it}},$$

which is equivalent to gross profit over sales.

COGS-share production-based markups (firm-FE Cobb–Douglas). We construct a production-based markup series that treats COGS as the flexible variable input under a Cobb–Douglas revenue production function with firm fixed effects. For each industry g, pooling all years, we estimate

$$\ln \mathrm{Sales}^{\$}_{it} = \beta_K \ln K_{it} + \beta_L \ln L_{it} + \theta_g \ln \mathrm{COGS}^{\$}_{it} + \alpha_i + u_{it},$$

where α_i is a firm fixed effect and θ_g is interpreted as the output elasticity with respect to COGS. Estimation is implemented by demeaning all logged variables within firm (within-industry) and running OLS on the demeaned system. We require a minimum industry sample size (baseline: 100 firm-year observations) to estimate θ_g; for smaller industries, θ_g and the resulting markups are left missing.

Let $s^{\mathrm{COGS}}_{it} = \mathrm{COGS}^{\$}_{it}/\mathrm{Sales}^{\$}_{it}$ denote the COGS revenue share. The implied markup is

$$\mu_{it} = \frac{\theta_g}{s^{\mathrm{COGS}}_{it}}.$$

We winsorize each markup series at the 1st and 99th percentiles (in levels) prior to aggregation and regression analysis. When studying $\ln(\mu_{it})$, we restrict to observations with $\mu_{it} > 0$; for event-study plots in levels and logs, we use a common positive-markup sample to keep composition fixed across the two transformations.

Finance & Insurance variants. Because interest expense can be an economically meaningful cost component for Finance & Insurance firms (NAICS 52), we consider alternative treatments of NAICS 52 when forming the flexible-input share. Our baseline measure uses $s_{it}^{\text{COGS}} = \text{COGS}_{it}^{\$}/\text{Sales}_{it}^{\$}$. As a robustness variant, we construct a Finance & Insurance adjustment that augments variable-input expenditure by adding interest expense only when it is observed:

$$s_{it}^{\text{COGS}} = \frac{\text{COGS}_{it}^{\$} + \text{XINT}_{it}^{\$}}{\text{Sales}_{it}^{\$}} \quad \text{for NAICS 52 firms when } \text{XINT}_{it}^{\$} \text{ is observed,}$$

and otherwise we fall back to the baseline share $\text{COGS}_{it}^{\$}/\text{Sales}_{it}^{\$}$ for NAICS 52 firms with missing XINT. We also report a Finance & Insurance exclusion variant that drops NAICS 52 firms entirely.

D.3 Factor shares and labor-share measures

Several outcomes in the paper require measures of factor shares and labor shares. Because Compustat does not report a comprehensive wage bill for most firms, we combine Compustat accounting data with micro-based labor aggregates from Revelio Labs.

Labor bill (wage bill). Let Users_{it} denote the Revelio firm–year headcount aggregate and let $\widetilde{w}_{it}$ denote the firm–year median annual salary. When available, we proxy the firm–year wage bill by

$$W_{it}^{\text{Rev}} = \text{Users}_{it} \times \widetilde{w}_{it}.$$

When Revelio-based wages are unavailable, we use an accounting fallback that treats a constant fraction of SG&A as labor expense:

$$W_{it}^{\text{SGA}} = 0.70 \times \text{XSGA}_{it}.$$

Our baseline wage-bill proxy is

$$W_{it} = \begin{cases} W_{it}^{\text{Rev}}, & \text{if } W_{it}^{\text{Rev}} \text{ is observed,} \\ W_{it}^{\text{SGA}}, & \text{otherwise.} \end{cases}$$

Materials and value added. We proxy intermediate inputs using COGS net of the wage bill when this net amount is positive, and otherwise fall back to COGS:

$$M_{it} = \begin{cases} \text{COGS}_{it} - W_{it}, & \text{if } \text{COGS}_{it} - W_{it} > 0, \\ \text{COGS}_{it}, & \text{otherwise.} \end{cases}$$

Value added is then

$$VA_{it} = \text{Sales}_{it} - M_{it},$$

defined only when $\text{Sales}_{it} > 0$ and $VA_{it} > 0$.

Capital cost proxy. We proxy the user cost of capital using depreciation plus the economic amortization of intangible capital (when constructed):

$$C_{it}^{K} = \text{DP}_{it} + \text{Amort}_{it}^{\text{intan}},$$

where $\text{Amort}_{it}^{\text{intan}}$ is the economic amortization defined in Appendix D.1. When $\text{Amort}_{it}^{\text{intan}}$ is not available, we set it to zero in this construction.

Value-added factor shares (α^{VA}). Define raw value-added shares

$$s_{it}^{K} = \frac{C_{it}^{K}}{VA_{it}}, \qquad s_{it}^{L} = \frac{W_{it}}{VA_{it}}.$$

Because VA_{it} includes operating surplus, $s_{it}^K + s_{it}^L$ need not equal one. We therefore renormalize:

$$\alpha_{K,it}^{VA} = \frac{s_{it}^K}{s_{it}^K + s_{it}^L}, \qquad \alpha_{L,it}^{VA} = \frac{s_{it}^L}{s_{it}^K + s_{it}^L}.$$

We trim extreme values by restricting $\alpha_{K,it}^{VA}, \alpha_{L,it}^{VA} \in [0.01, 0.99]$.

Gross-output (GO3) shares (α^{GO3}). For the gross-output specification, we first compute a materials share in revenue, $s_{it}^M = M_{it}/\text{Sales}_{it}$, and set the materials elasticity to the industry–year median:

$$\alpha_{M,it}^{GO3} = \text{median}_{(g,t)} \left(s_{it}^M \right),$$

clipped to a conservative range. We then allocate the remaining share, $1 - \alpha_{M,it}^{GO3}$, between capital and labor using the value-added shares:

$$\alpha_{K,it}^{GO3} = (1 - \alpha_{M,it}^{GO3})\alpha_{K,it}^{VA}, \qquad \alpha_{L,it}^{GO3} = (1 - \alpha_{M,it}^{GO3})\alpha_{L,it}^{VA},$$

and finally renormalize ($\alpha_{K,it}^{GO3} + \alpha_{L,it}^{GO3} + \alpha_{M,it}^{GO3} = 1$).

Industry–year pooling for stability. To stabilize event-time and dispersion analyses, we use industry–year pooling for $\alpha_{K,it}^{VA}$ and $\alpha_{L,it}^{VA}$ when industry–year cell sizes are sufficiently large. Concretely, we replace firm-level raw shares with industry–year means when the cell contains at least a minimum number of observations, and otherwise leave the cell missing.

Alternative Labor-share measure. In addition to the factor-share objects above, we report complementary labor-share measures designed to be transparent about the role of wage-bill imputation. Define a value-added proxy that subtracts COGS and the Peters-style organizational-investment component of SG&A:

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

defined only when $VA_{it}^{A} > 0$. We then construct:

$$\text{Labor Share}_{it}^{Alt} = \frac{W_{it}^{\text{Alt}}}{VA_{it}^{A}},$$

where W_{it}^{Alt} uses XLR_{it} when available and otherwise $0.70 \times \text{XSGA}_{it}$.

E Who adopts AI, and when do they adopt?

E.1 Adoption hazard: industry fixed effects

Figure 34 reports a SHAP summary for the XGBoost discrete-time hazard classifier (cross-validated mean AUC = 0.872). The lagged leave-one-out diffusion state $S_{g,t-1}^{-i}$ and firm size (log market capitalization) are the dominant predictors of first AI adoption. Secondary contributions come from liquidity (cash-to-assets) and intangible intensity (intangibles-to-assets). Overall, the ranking is consistent with a frontier/capabilities selection mechanism: larger firms with complementary resources are more likely to adopt, and adoption occurs earlier when diffusion within the firm's industry is already advanced.

Table 12 lists the industries with the highest and lowest adoption odds in the ridge-logit hazard model, based on the estimated industry fixed effects (relative to the omitted industry group). These odds ratios summarize baseline differences in adoption propensity after absorbing year effects and controlling for standardized firm covariates; they should be interpreted as descriptive correlates rather than causal effects.

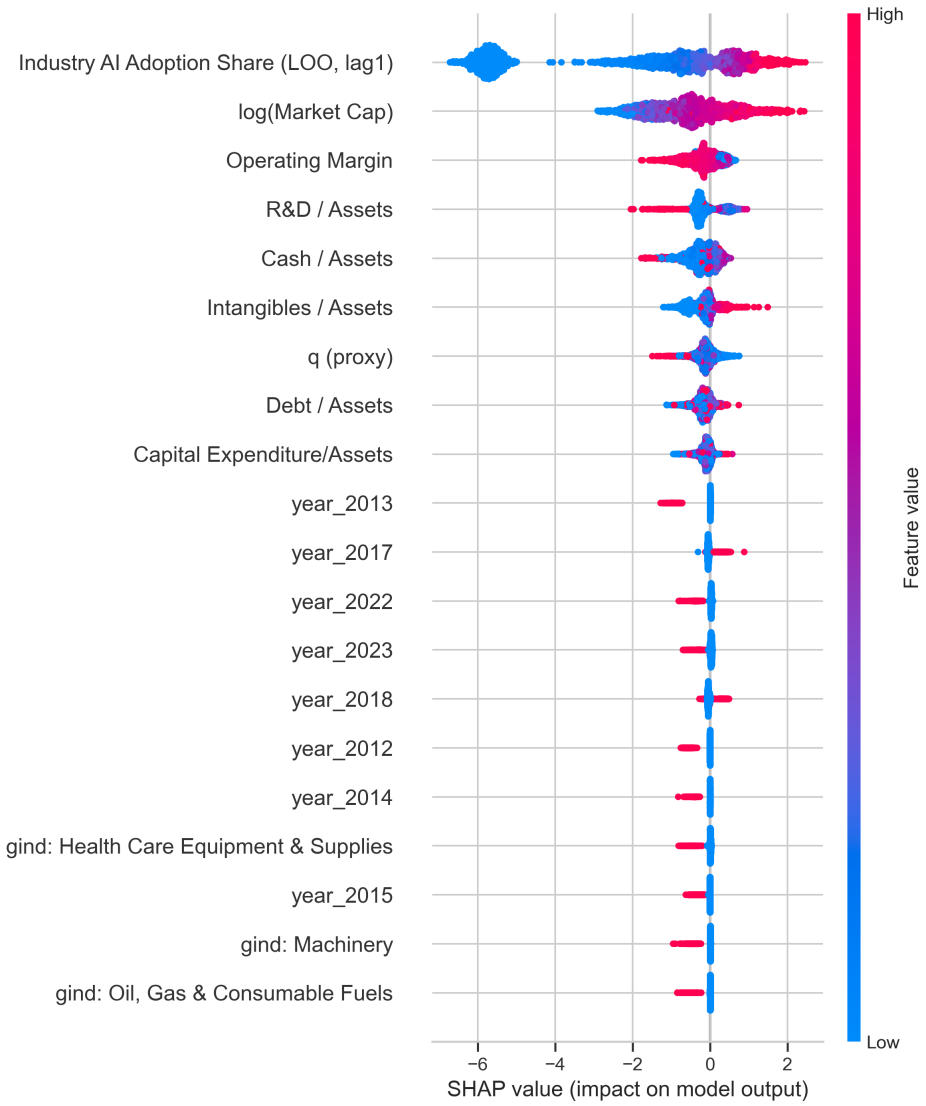


Figure 34: SHAP summary for the adoption hazard model (XGBoost). Higher industry diffusion $S^{-i}_{g,t-1}$ and larger firm size are the dominant predictors of first AI adoption, with additional contributions from cash holdings and intangible intensity.

Table 12: AI adoption hazard: industry fixed-effect odds ratios (top/bottom 10)

Top 10 industries (hazard model)		
GICS	Industry	OR
201040	Electrical Equipment	1.773
203020	Airlines	1.650
255040	Specialty Retail	1.581
452030	Electronic Equipment, Instruments & Components	1.434
402010	Diversified Financial Services	1.371
253010	Hotels, Restaurants & Leisure	1.340
551050	Independent Power and Renewable Electricity Producers	1.322
251020	Automobiles	1.311
402020	Consumer Finance	1.289
203040	Road & Rail	1.287
Bottom 10 industries (hazard model)		
401010	Banks	0.803
303020	Personal Products	0.752
352010	Biotechnology	0.704
302030	Tobacco	0.692
203030	Marine	0.667
302010	Beverages	0.540
402040	Mortgage REITs / NA	0.464
551020	Gas Utilities	0.447
151040	Metals & Mining	0.439
352020	Pharmaceuticals	0.432

E.2 Ever vs. never adopters: industry and cohort fixed effects

Table 13 reports the top and bottom industries by the estimated industry fixed effects in the ever-vs-never ridge logit, controlling for first-observed-year cohort fixed effects. As above, the odds ratios reflect baseline differences in adoption propensity relative to the omitted industry group.

Table 13: Ever vs. never (AI): industry fixed-effect odds ratios (top/bottom 10)

Top 10 industries (ever vs. never)		
GICS	Industry	OR
202020	Professional Services	2.586
201040	Electrical Equipment	2.162
451020	IT Services	1.997
452030	Electronic Equipment, Instruments & Components	1.855
502010	Media (NA in raw output)	1.605
402020	Consumer Finance	1.593
403010	Insurance	1.446
452020	Technology Hardware, Storage & Peripherals	1.400
452010	Communications Equipment	1.391
352030	Life Sciences Tools & Services	1.370
Bottom 10 industries (ever vs. never)		
302020	Food Products	0.676
401010	Banks	0.673
352010	Biotechnology	0.619
151010	Chemicals	0.598
101020	Oil, Gas & Consumable Fuels	0.482
303020	Personal Products	0.420
402040	Mortgage REITs / NA	0.273
352020	Pharmaceuticals	0.193
151040	Metals & Mining	0.086

ML prediction. We also train a gradient-boosted decision tree classifier (XGBoost) to predict whether a firm will ever adopt AI. The model achieves a holdout AUC of 0.972 and a cross-validated mean AUC of 0.981. Figure 35 plots SHAP values summarizing each feature's contribution to predicted adoption probabilities. As in the hazard model, $S^{-i}_{g,t-1}$ and firm size dominate, followed by cash-to-assets and other firm characteristics.

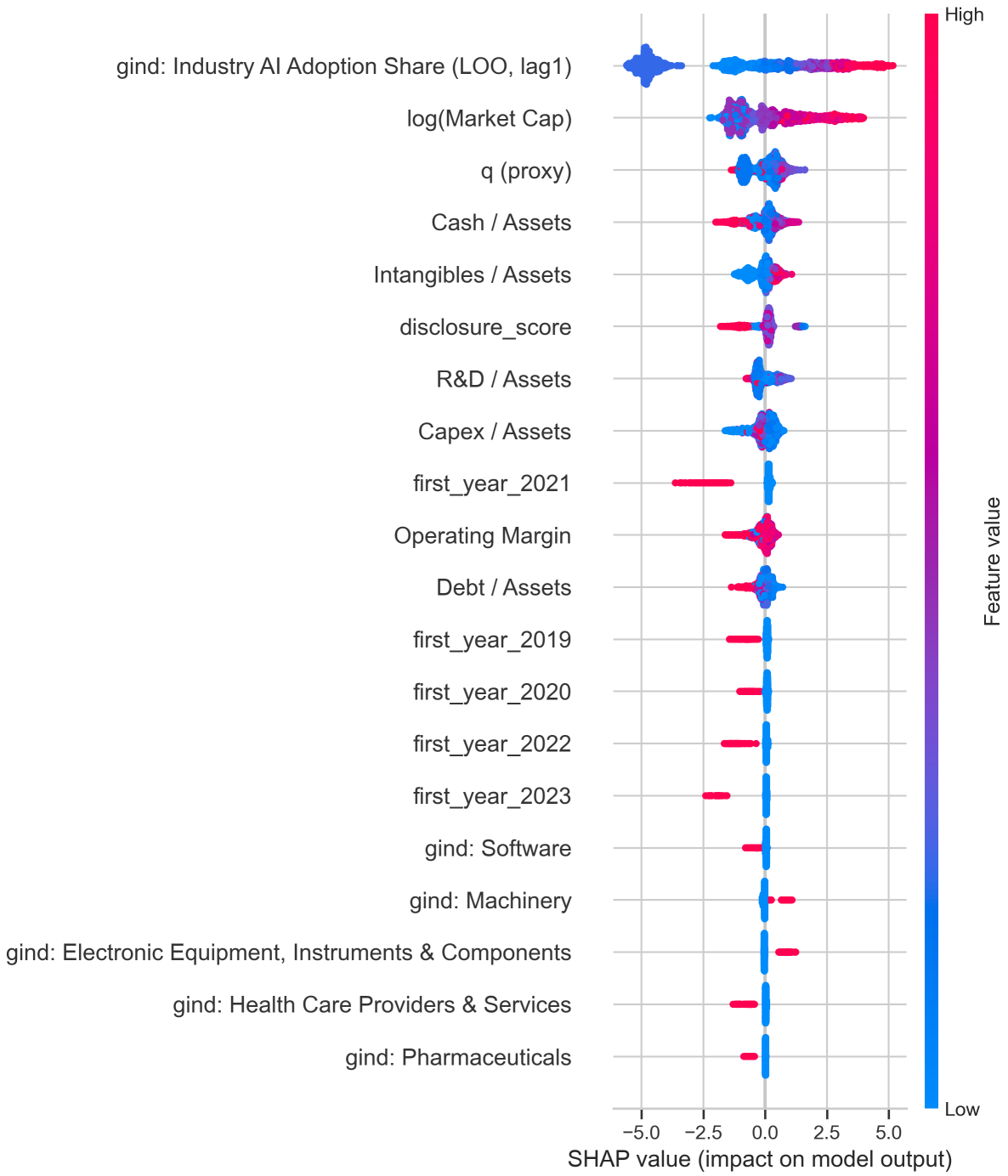


Figure 35: SHAP summary (beeswarm) for the ever-vs-never AI adoption classifier (XGBoost). Each point is a firm observation; the x-axis reports the SHAP value (feature contribution to the model output), where positive values increase the predicted probability of ever adopting AI and negative values decrease it. Colors indicate feature values (red = high, blue = low). The dominant predictors are the lagged leave-one-out industry diffusion measure $S^{-i}_{g,t-1}$ and firm size (log market cap): higher diffusion and larger firms strongly shift predictions toward ever adoption. Secondary contributors include valuation (q proxy), cash holdings, intangible intensity, and investment variables, while cohort (first-observed-year) and industry indicators capture residual differences in baseline adoption propensity.

F Markup levels (robustness)

This appendix reports event-time dynamics for markup levels for the three production-based constructions used as pricing-power diagnostics. Because markup levels are harder to interpret in absolute units and are more sensitive to outliers and scale, the main text emphasizes log markups. The level plots here serve as a robustness check that the qualitative conclusions are not driven by the log transformation. Figure 36 shows that, in levels as well, there is no robust evidence of a sustained post-adoption increase in markups under (a) the full-sample COGS-share construction, (b) the Finance & Insurance adjustment that augments COGS by interest expense for NAICS 52 firms when interest expense is observed prior to forming the variable-input share, or (c) the Finance & Insurance exclusion variant. Post-adoption movements are small relative to sampling uncertainty and are not stable across horizons or adoption vintages. These patterns support the interpretation in Section 4.2.3 that the paper's medium-run profitability dynamics are unlikely to be driven primarily by rising pricing power.

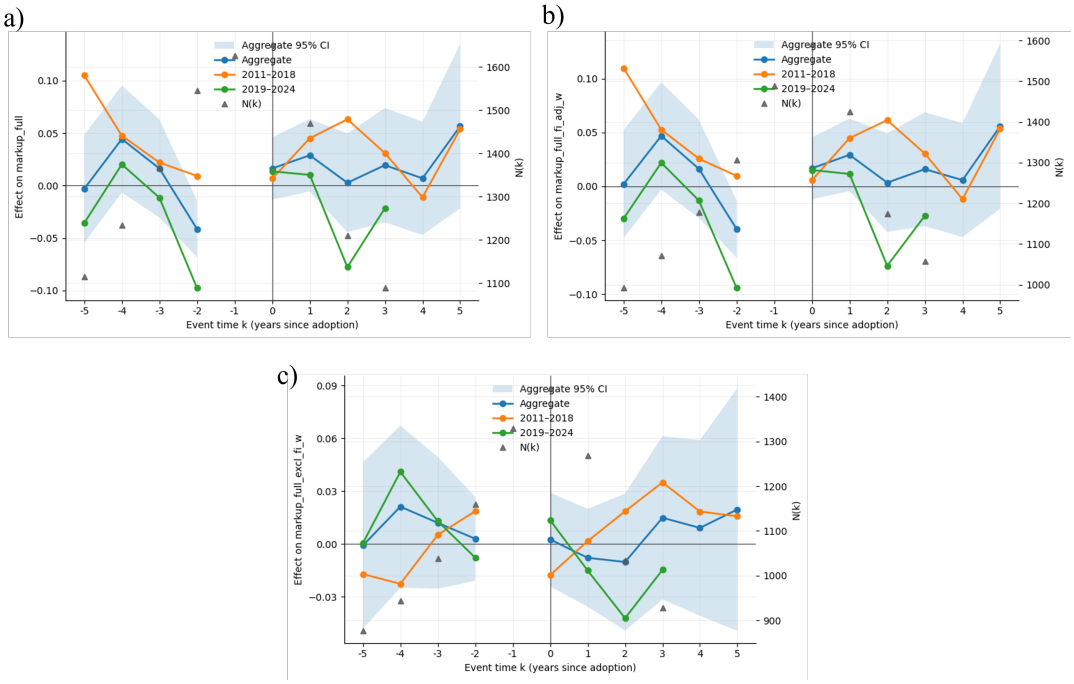


Figure 36: Markup levels around AI adoption for three constructions: (a) full-sample COGS-share markups, (b) a Finance & Insurance adjusted markup, and (c) a Finance & Insurance exclusion variant. The figure plots aggregate Sun–Abraham IW coefficients with 95% confidence intervals, along with adoption-vintage overlays.

G Pilot-only adopters

This appendix reports event-study dynamics for pilot-only firms using the same score rubric and event-study setup as Section 4.4. Let $\text{AID}_{it} \in \{0, 0.5, 1\}$ denote the discrete AI Adoption score from transcript coding, where 0.5 corresponds to pilot/announcement evidence and 1 corresponds to confirmed deployment. Define

$$G_i^{0.5} = \min\{t : \text{AID}_{it} \geq 0.5\}$$

as the first fiscal year in which firm i crosses the pilot/announcement threshold. We classify a firm as pilot-only if it never reaches confirmed deployment within the sample window, i.e.,

$$\max_t \text{AID}_{it} < 1.$$

We estimate Sun–Abraham event studies around $G_i^{0.5}$ for the pilot-only sample, using the same never-adopter comparison group and restricting event-time support to $k \in [-5, +5]$ to keep support fixed. The resulting event-time paths show comparatively weak and less persistent movements in valuation and operating outcomes for pilot-only firms. This pattern is consistent with the interpretation that the baseline effects are primarily associated with confirmed deployment rather than pilot/announcement activity alone.

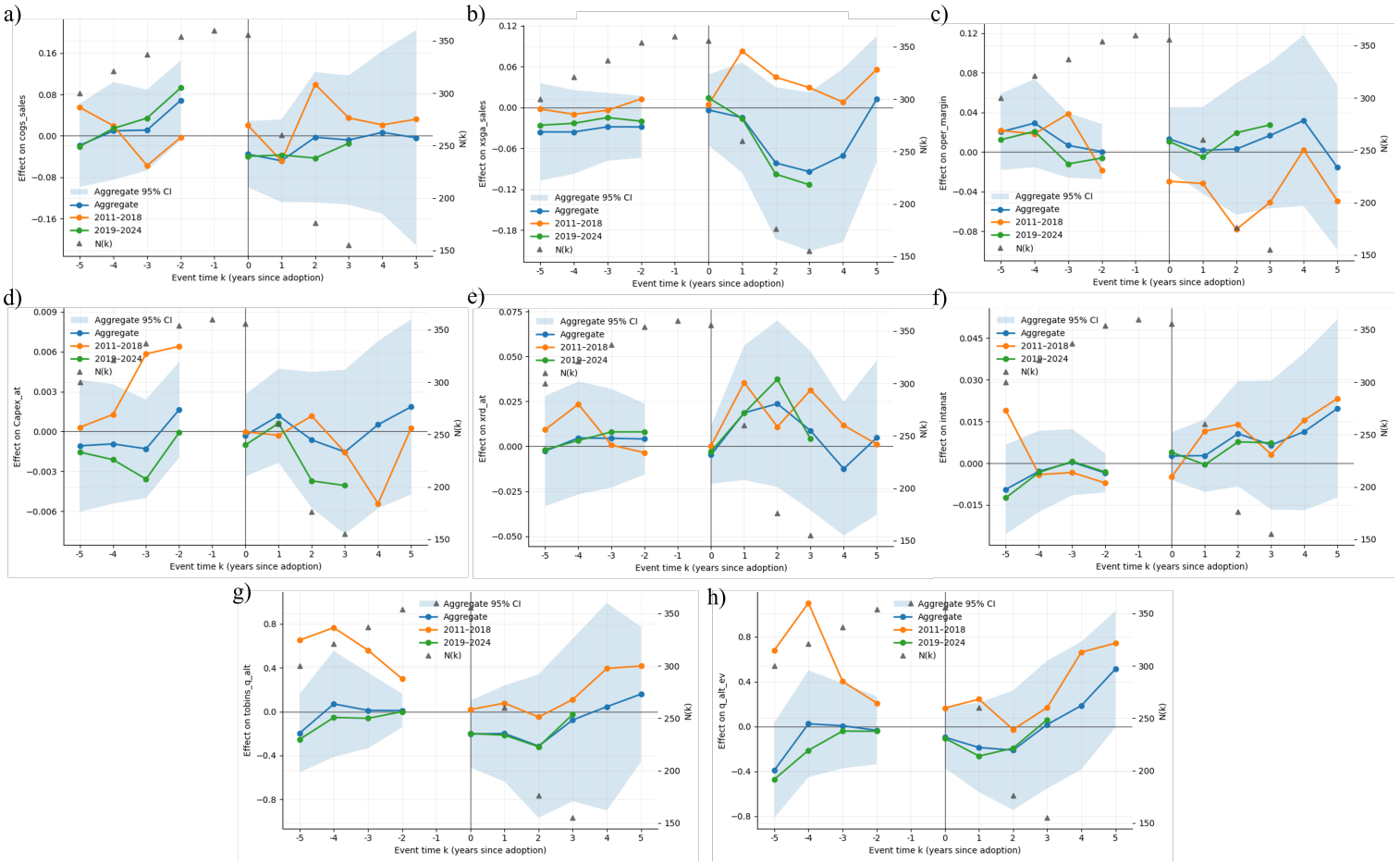


Figure 37: Pilot-only adopters (event time $G_i^{0.5}$). The sample consists of firms whose AI Adoption score reaches the pilot/announcement threshold (AID ≥ 0.5) but never reaches confirmed deployment (AID $= 1$) within the sample. The figure plots Sun–Abraham event-study coefficients with 95% confidence intervals; triangles denote the number of observations by event time. Relative to the main deployment sample, pilot-only firms exhibit weaker and less persistent changes across outcomes, consistent with deployment (rather than announcement) driving the primary effects.

H IV identification, diagnostics, and robustness (early/late specification)

This appendix reports identification diagnostics and robustness checks for the vintage-heterogeneous 2SLS specification in Table 3, which distinguishes early (2011–2018) and late (2019–2024) AI adoption under a diffusion-share instrument. Unless otherwise noted, standard errors are clustered two ways by firm and year.

H.1 Setup and excluded instruments

The second-stage equation is

$$Y_{it} \;=\; \beta_E\, AI^{E}_{it} \;+\; \beta_L\, AI^{L}_{it} \;+\; \mathbf{X}'_{it-1}\gamma \;+\; \alpha_i \;+\; \lambda_t \;+\; \varepsilon_{it}, \tag{H.1}$$

where AI^{E}_{it} and AI^{L}_{it} are the early- and late-adoption components of the absorbing AI adoption state, $\mathbf{X}_{it-1}$ are baseline lagged controls (lagged log market cap and lagged R&D/Assets), and α_i and λ_t are firm and year fixed effects.

Let $Z_{g(i),t-1}$ denote the lagged leave-one-out industry diffusion share for firm i's industry $g(i)$. The excluded instruments are the vintage-interacted shifters

$$Z^{E}_{it} \equiv Z_{g(i),t-1} \cdot \mathbf{1}\{G_i < 2019\}, \qquad Z^{L}_{it} \equiv Z_{g(i),t-1} \cdot \mathbf{1}\{G_i \geq 2019\},$$

where G_i is firm i's first strict adoption year (undefined for never adopters).

H.2 First-stage relevance

For each endogenous regressor AI^{k}_{it} with $k \in \{E, L\}$, we estimate

$$AI^{k}_{it} \;=\; \Pi_k\, Z^{E}_{it} \;+\; \Lambda_k\, Z^{L}_{it} \;+\; \mathbf{X}'_{it-1}\delta_k \;+\; \alpha_i \;+\; \lambda_t \;+\; u^{k}_{it}. \tag{H.2}$$

We summarize relevance using (i) the partial R^2 of the excluded instruments relative to a specification with only controls and fixed effects, and (ii) a clustered Wald test of the

joint significance of the excluded instruments. Table 14 reports these diagnostics; the excluded instruments are strongly predictive for both AI_{it}^{E} and AI_{it}^{L}.

Table 14: First-stage relevance diagnostics

Endogenous regressor	N	Partial R^2	Wald stat	p-value	df
AI_early	19017	0.3683	47.0212	0.0000	2
AI_late	19017	0.3717	189.7776	0.0000	2

Notes: This table reports first-stage diagnostics estimated once on the fixed sample defined by non-missing endogenous variables, excluded instruments, and controls. "Partial R^2" is the partial R^2 of the excluded instruments relative to a specification with controls, firm fixed effects, and year fixed effects. "Wald" reports a clustered Wald test of the joint significance of the excluded instruments. Standard errors in first-stage regressions are clustered two-way by firm and year.

H.3 Endogeneity and weak-IV-robust inference (by outcome)

Because identification strength and endogeneity concerns can vary by outcome, we report two outcome-by-outcome diagnostics.

Endogeneity test (Wu–Hausman). We implement a Wu–Hausman endogeneity test via a control-function approach with firm and year fixed effects: we augment the outcome equation with the first-stage residual(s) for the endogenous regressor(s) and test their joint significance. Failure to reject suggests that endogeneity is not strongly detected for that outcome under our specification; rejection suggests that IV is warranted.

Weak-IV-robust inference (Anderson–Rubin style). We also report an Anderson–Rubin (AR)-style test based on the reduced-form regression of Y_{it} on the excluded instruments, controls, and fixed effects. The AR test remains valid under weak identification and provides a conservative check of whether the excluded instruments have predictive content for the outcome through the endogenous channel.

Table 15 reports both the Wu–Hausman and AR-style diagnostics. Across the key outcomes emphasized in the paper (scale, cost shares, intangible accumulation, and financing), the diagnostics typically support the relevance of the excluded instruments and are consistent with the use of IV. For some outcomes, neither test indicates strong evidence of an IV-detectable channel in our specification, consistent with the corresponding 2SLS

estimates being economically small and statistically imprecise. The early/late specification uses two excluded instruments for two endogenous regressors and is therefore exactly identified; overidentification tests (e.g., Hansen's J) are not applicable.

Table 15: IV diagnostics by outcome

Outcome	**Wu–Hausman**	**Anderson–Rubin (AR)**	N
Sales	6.1542**	8.2282**	19017
	(0.0461)	(0.0163)	
Operating margin	7.2777**	7.0400**	19017
	(0.0263)	(0.0296)	
COGS / Sales	13.5466***	13.3320***	19017
	(0.0011)	(0.0013)	
SG&A / Sales	12.0172***	11.3143***	16626
	(0.0025)	(0.0035)	
R&D / Assets	1.5328	4.4767	18876
	(0.4647)	(0.1066)	
Tobin's q_{alt}	3.8025	11.8702***	19013
	(0.1494)	(0.0026)	
EV-based q_{alt}^{EV}	9.7424***	21.0712***	19017
	(0.0077)	(0.0000)	
Debt / Assets	6.5041**	6.1094**	18928
	(0.0387)	(0.0471)	
Cash / Assets	10.6233***	13.6729***	19017
	(0.0049)	(0.0011)	
Capex / Assets	0.1337	1.0173	18989
	(0.9353)	(0.6013)	
Intangibles / Assets	7.8089**	16.5160***	18829
	(0.0202)	(0.0003)	
Intangible capital K^{intan}	7.2602**	7.2294**	16941
	(0.0265)	(0.0269)	
ln(Markup) (full)	9.8648***	10.8952***	18133
	(0.0072)	(0.0043)	
ln(Markup) (excl. financials)	8.5650**	9.5953***	17861
	(0.0138)	(0.0082)	
ln(Markup) (FI-adjusted)	9.4170***	10.8651***	18133
	(0.0090)	(0.0044)	
Operating margin (adj)	0.8807	0.9786	16941
	(0.6438)	(0.6130)	
Labor share (Alt)	2.3833	1.4647	15103
	(0.3037)	(0.4808)	
Capital share α_K^{VA}	4.2057	3.8825	19012
	(0.1221)	(0.1435)	
Labor share α_L^{VA}	4.2057	3.8825	19012
	(0.1221)	(0.1435)	
Capital share α_K^{GO3}	6.2280**	6.4985**	19012
	(0.0444)	(0.0388)	
Labor share α_L^{GO3}	5.8129*	6.1852**	19012
	(0.0547)	(0.0454)	

Notes: Each row reports diagnostics from the corresponding 2SLS regression with firm and year fixed effects and baseline lagged controls. Entries report the test statistic, with p-values in parentheses. Significance: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

H.4 Robustness to diffusion timing and industry-year shocks

A common concern in diffusion-share designs is that industry-level adoption dynamics may proxy for time-varying industry opportunity or industry-cycle timing. Our preferred instrument uses lagged leave-one-out diffusion, $Z_{g(i),t-1}$, which is predetermined relative to outcomes in year t.

A natural robustness approach would be to absorb industry-year shocks using industry×year fixed effects; however, because the diffusion shifter is defined at the industry-year level, industry×year fixed effects absorb the excluded instrument by construction and render the 2SLS model unidentified. We therefore use three feasible substitutes and report the resulting vintage-heterogeneous 2SLS estimates.

(i) Alternative diffusion timing: contemporaneous LOO diffusion. Appendix Table 16 replaces the lagged diffusion share with contemporaneous LOO diffusion. While this shifter is less predetermined (and therefore not preferred), it provides a sensitivity check on whether the baseline results are mechanically driven by the one-year lag choice. The key patterns match the baseline in Table 3: early adopters expand scale and exhibit higher COGS/Sales and SG&A/Sales with lower operating margins, while both vintages show increased intangible intensity and a shift in factor shares away from labor.

Table 16: Vintage-heterogeneous 2SLS effects of AI adoption on firm outcomes: contemporaneous diffusion share with two-way clustered standard errors

	Early AI (2011–2018)	**Late AI (2019–2024)**	N
Scale and cost structure			
Sales (USD millions)	11002.6***	2559.8*	19016
	(4243.3)	(1405.4)	
COGS / Sales	0.196***	0.098**	19016
	(0.057)	(0.046)	
SG&A / Sales	0.117**	-0.031	16625
	(0.047)	(0.056)	
Operating margin	-0.099***	-0.021	19016
	(0.033)	(0.031)	
Intangibles and investment mix			
Intangibles / Assets	0.057***	0.062***	18828
	(0.018)	(0.016)	
R&D / Assets	0.027**	0.005	18875
	(0.011)	(0.013)	
Capex / Assets	-0.003	-0.002	18988
	(0.003)	(0.002)	
Intangible capital K^{intan} (USD millions)	6880.9**	509.6**	16940
	(2767.3)	(238.9)	
Factor shares			
Capital share α_K^{VA}	0.116**	0.080**	19012
	(0.052)	(0.036)	
Labor share α_L^{VA}	-0.116**	-0.080**	19012
	(0.052)	(0.036)	
Capital share α_K^{GO3}	0.080***	0.048***	19012
	(0.027)	(0.016)	
Labor share α_L^{GO3}	-0.073***	-0.044***	19012
	(0.027)	(0.016)	
Valuation			
Tobin's q_{alt}	0.463**	-0.449*	19012
	(0.221)	(0.268)	
EV-based q_{alt}^{EV}	1.327***	-0.128	19016
	(0.312)	(0.447)	
Financing			
Debt / Assets (book leverage)	0.066***	0.023	18927
	(0.023)	(0.020)	
Cash / Assets	-0.068***	-0.042***	19016
	(0.017)	(0.014)	

Notes: Each row is a separate 2SLS regression with firm and year fixed effects and lagged controls (lagged log market cap and lagged R&D/Assets). Standard errors are clustered two ways by firm and year (reported in parentheses). Significance: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$. Rows shown as '–' were not produced in the provided robustness output.

(ii) Controlling for lagged industry-year conditions. Appendix Table 17 augments the baseline specification with lagged industry-year aggregates of slow-moving conditions (e.g., sales growth, operating margins, and R&D intensity). These controls directly address the concern that diffusion captures predictable industry trends. The resulting coefficients

are very close to the baseline estimates in Table 3 for the main mechanisms (cost shares, intangible accumulation, factor-share reallocation, and financing), indicating that the diffusion instrument is not simply proxying for lagged industry conditions.

Table 17: Vintage-heterogeneous 2SLS effects of AI adoption: controlling for lagged industry-year outcomes (two-way clustered SEs)

	Early AI (2011–2018)	**Late AI (2019–2024)**	N
Scale and cost structure			
Sales (USD millions)	10874.8**	2535.2*	19016
	(4227.4)	(1381.9)	
COGS / Sales	0.187***	0.089**	19016
	(0.056)	(0.043)	
SG&A / Sales	0.116**	-0.034	16625
	(0.045)	(0.055)	
Operating margin	-0.092***	-0.015	19016
	(0.031)	(0.030)	
Intangibles and investment mix			
Intangibles / Assets	0.056***	0.062***	18828
	(0.018)	(0.016)	
R&D / Assets	0.028**	0.005	18875
	(0.011)	(0.013)	
Capex / Assets	-0.003	-0.001	18988
	(0.003)	(0.002)	
Intangible capital K^{intan} (USD millions)	6910.5**	547.5**	16940
	(2755.3)	(226.4)	
Factor shares			
Capital share α_K^{VA}	0.129***	0.083***	19012
	(0.043)	(0.031)	
Labor share α_L^{VA}	-0.129***	-0.083***	19012
	(0.043)	(0.031)	
Capital share α_K^{GO3}	0.085***	0.049***	19012
	(0.025)	(0.015)	
Labor share α_L^{GO3}	-0.076***	-0.042***	19012
	(0.025)	(0.016)	
Valuation			
EV-based q_{alt}^{EV}	1.107***	-0.257	19016
	(0.293)	(0.420)	
Financing			
Debt / Assets (book leverage)	0.064***	0.022	18927
	(0.023)	(0.020)	
Cash / Assets	-0.070***	-0.043***	19016
	(0.017)	(0.015)	

Notes: Same as Appendix Table 16, but includes lagged industry-year aggregates as additional controls. Standard errors clustered two ways by firm and year. Significance: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$. Rows shown as '–' were not produced in the provided robustness output.

(iii) Coarser industry diffusion instrument. Appendix Table 18 reconstructs the diffusion share at a coarser industry aggregation. This reduces simultaneity concerns that can be more acute at highly granular industry definitions, at the cost of weaker

within-industry identifying variation. Consistent with reduced power, some late-vintage effects become less precisely estimated (e.g., scale and valuation), but the qualitative pattern for the paper's central mechanisms remains aligned with the baseline: adoption is associated with higher intangible intensity and a reallocation of factor shares away from labor, and early adopters display a cost-share-driven margin compression consistent with an expensed build-out phase.

Table 18: Vintage-heterogeneous 2SLS effects of AI adoption: coarser-industry diffusion instrument (two-way clustered SEs)

	Early AI (2011–2018)	**Late AI (2019–2024)**	*N*
Scale and cost structure			
Sales (USD millions)	8610.0**	886.2	19247
	(3401.5)	(780.3)	
COGS / Sales	0.204***	0.094**	19247
	(0.063)	(0.048)	
SG&A / Sales	0.120**	-0.015	16853
	(0.050)	(0.059)	
Operating margin	-0.088**	-0.016	19247
	(0.034)	(0.033)	
Intangibles and investment mix			
Intangibles / Assets	0.065***	0.060***	19059
	(0.018)	(0.015)	
Capex / Assets	-0.002	-0.001	19219
	(0.003)	(0.003)	
Intangible capital K^{intan} (USD millions)	6215.0***	442.9*	17168
	(2258.9)	(268.4)	
Factor shares			
Capital share α_K^{VA}	0.073	0.041	19012
	(0.046)	(0.027)	
Labor share α_L^{VA}	-0.073	-0.041	19012
	(0.046)	(0.027)	
Capital share α_K^{GO3}	0.052*	0.025*	19012
	(0.030)	(0.015)	
Labor share α_L^{GO3}	-0.051*	-0.024	19012
	(0.029)	(0.015)	
Valuation			
EV-based q_{alt}^{EV}	1.719***	0.084	19247
	(0.352)	(0.446)	
Financing			
Debt / Assets (book leverage)	0.043*	0.014	19158
	(0.023)	(0.019)	
Cash / Assets	-0.071***	-0.036**	19247
	(0.019)	(0.014)	

Notes: Same as Appendix Table 16, but the diffusion-share instrument is constructed at a coarser industry aggregation and interacted with early/late adoption indicators. Standard errors clustered two ways by firm and year. Significance: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$. Rows shown as '–' were not produced in the provided robustness output.

Across these three feasibility-robust substitutes, the baseline interpretation is unchanged: IV estimates continue to point to post-adoption movements in cost shares, intangible investment/accumulation, and factor shares (away from labor), rather than an increase in pricing power.

I Additional diffusion robustness

A mechanical concern in diffusion-share designs is that diffusion may proxy for time-varying industry opportunity (or industry-cycle timing), or be intertwined with same-year industry outcomes. Two guardrails are especially direct in our setting. First, we assign groups using a diffusion state measured strictly prior to the adoption year. In the main text, the predetermined diffusion-at-adoption measure is $S_i \equiv S^{-i}_{g(i),G_i-1}$, the leave-one-out industry adoption share in the year before firm i adopts. Second, we verify that the qualitative ranking is not sensitive to whether diffusion is measured one versus two years before adoption, and we also report a contemporaneous (adoption-year) split as a diagnostic.

Figure 38 repeats the first-mover/follower IW event studies using diffusion measured two years before adoption, $S^{-i}_{g(i),G_i-2}$. The qualitative ranking is unchanged: first movers revalue earlier and by more than followers. Figure 39 reports the contemporaneous classification based on $S^{-i}_{g(i),G_i}$. This split is directionally similar but noisier, consistent with the fact that contemporaneous industry-year diffusion is more exposed to same-year measurement and can yield smaller effective support for the low-diffusion group in some industries.

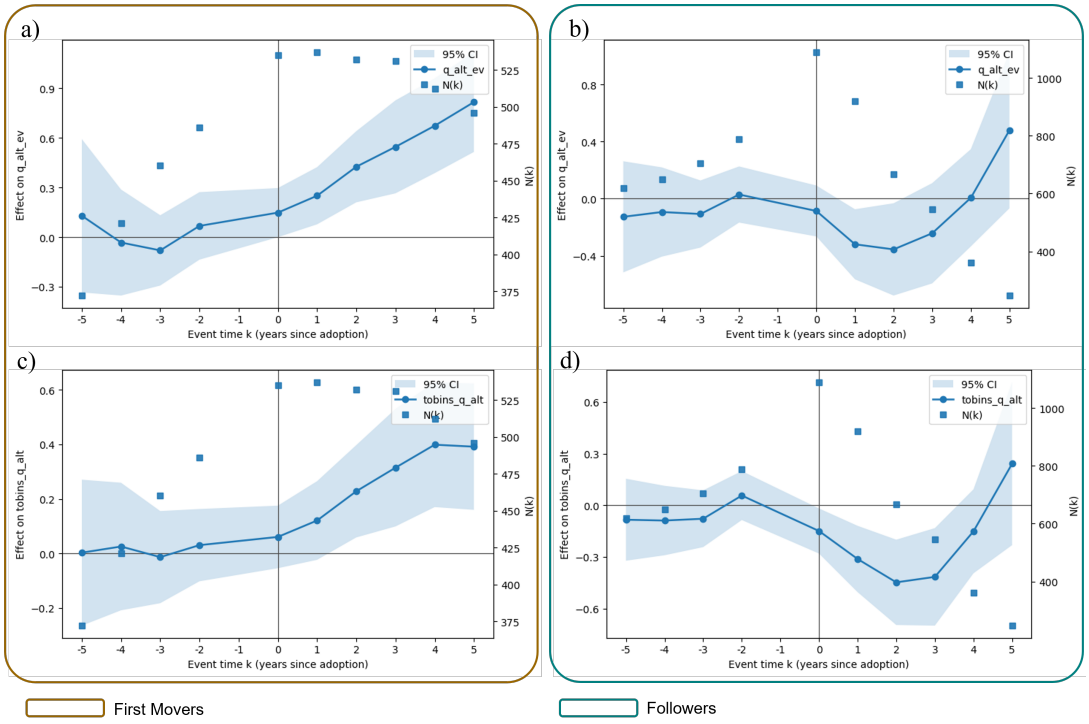


Figure 38: Diffusion at adoption and valuation: robustness using $S^{-i}_{g(i),G_i-2}$. Same design as Figure 11, except first movers vs. followers are defined using the diffusion state measured two years prior to adoption, $S^{-i}_{g(i),G_i-2}$. Left panels show first movers; right panels show followers.

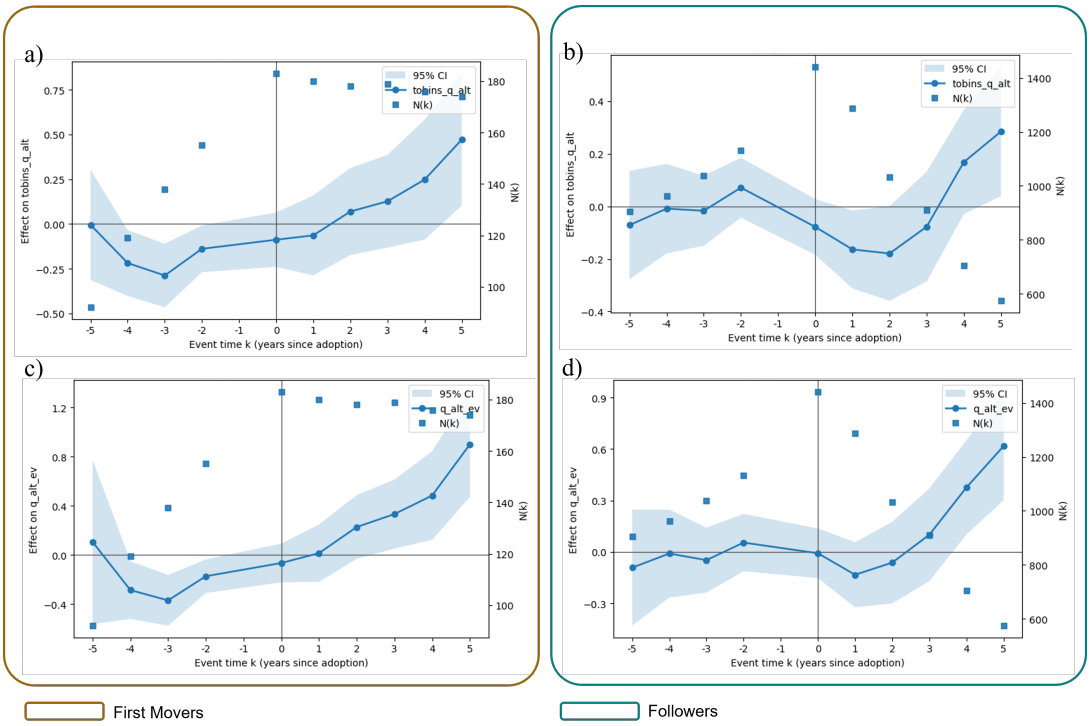


Figure 39: Diffusion at adoption and valuation: contemporaneous split using $S^{-i}_{g(i),G_i}$. First movers vs. followers are defined using the leave-one-out industry adoption share in the adoption year. Left panels show first movers; right panels show followers.

J Additional heterogeneity results

Figure 40 reports additional heterogeneity by vendor role for the production-function share objects. Vendor-type adopters exhibit a more negative post-adoption drift in the materials-share proxy α_M^{GO3} at longer horizons, whereas non-vendors remain closer to zero. For the labor share proxy α_L^{VA}, both vendor and non-vendor groups decline after adoption, but the decline is smaller for vendors. Because these objects are estimated shares and event-time support falls at the tails, longer-horizon movements should be interpreted cautiously (triangles report $N(k)$).

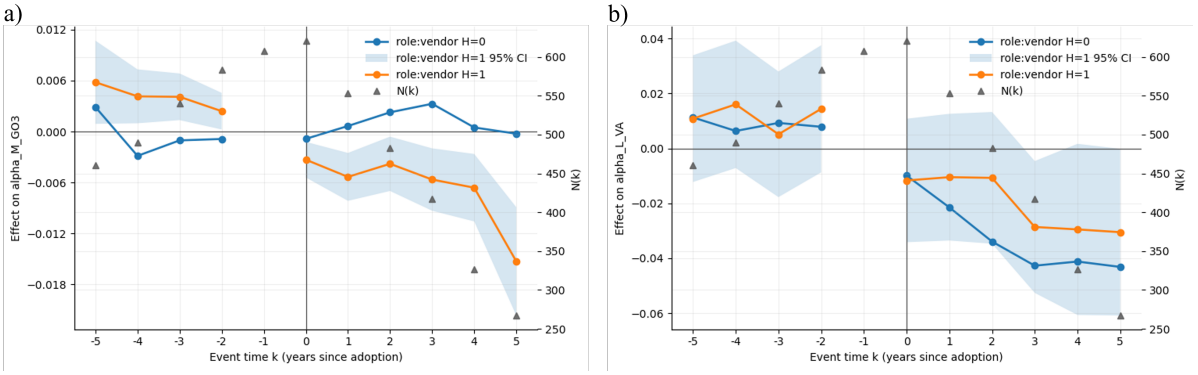


Figure 40: Additional heterogeneity by vendor role. **(a)** Materials-share proxy α_M^{GO3}. **(b)** Labor-share proxy α_L^{VA}. Classified-only sample ($ANY_ROLE = 1$). Triangles report $N(k)$.

K Additional mechanism results by intensity

Figure 41 reports additional outcomes for the intensity-based mechanism exercise. Valuation (Tobin's q) rises after adoption and exhibits a clear gradient with integration intensity under both intensity measures. Operating margins show mid-horizon improvements that are stronger at higher intensity, with a late-horizon reversal where sampling uncertainty is larger due to smaller $N(k)$. R&D intensity increases after adoption but is lower along the high-intensity path at longer horizons, consistent with intensive adopters shifting investment from expensed R&D toward broader intangible accumulation. Finally, higher integration intensity is associated with larger reductions in the materials share proxy α_M^{GO3}. Results are similar under the alternative (flow vs. stock) intensity definition.

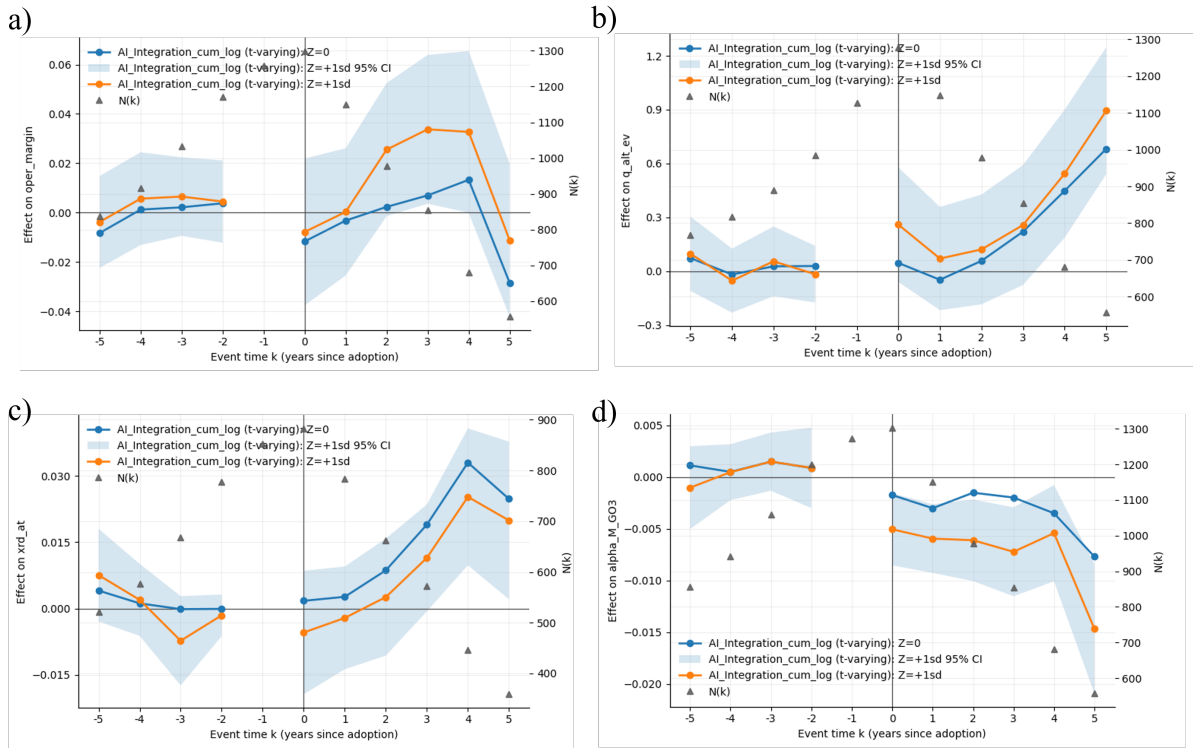


Figure 41: Additional mechanisms by time-varying integration intensity. (a) Alternative Tobin's q using cumulative integration. (b) Alternative Tobin's q using integration level. (c) Operating margin. (d) R&D/assets. (e) Materials-share proxy α_M^{GO3}. Triangles report $N(k)$.

L Additional evidence on capability build-out and rent compression

L.1 Predetermined split: first movers vs. followers

This subsection complements the calendar-vintage heterogeneity by using a predetermined diffusion split at the firm's adoption date (as in Section 4.3). Firms are classified as first movers vs. followers based on the leave-one-out industry diffusion state measured strictly prior to adoption, $S^{-i}_{g(i),G_i-1}$, ensuring that group assignment is predetermined with respect to adoption-year outcomes.

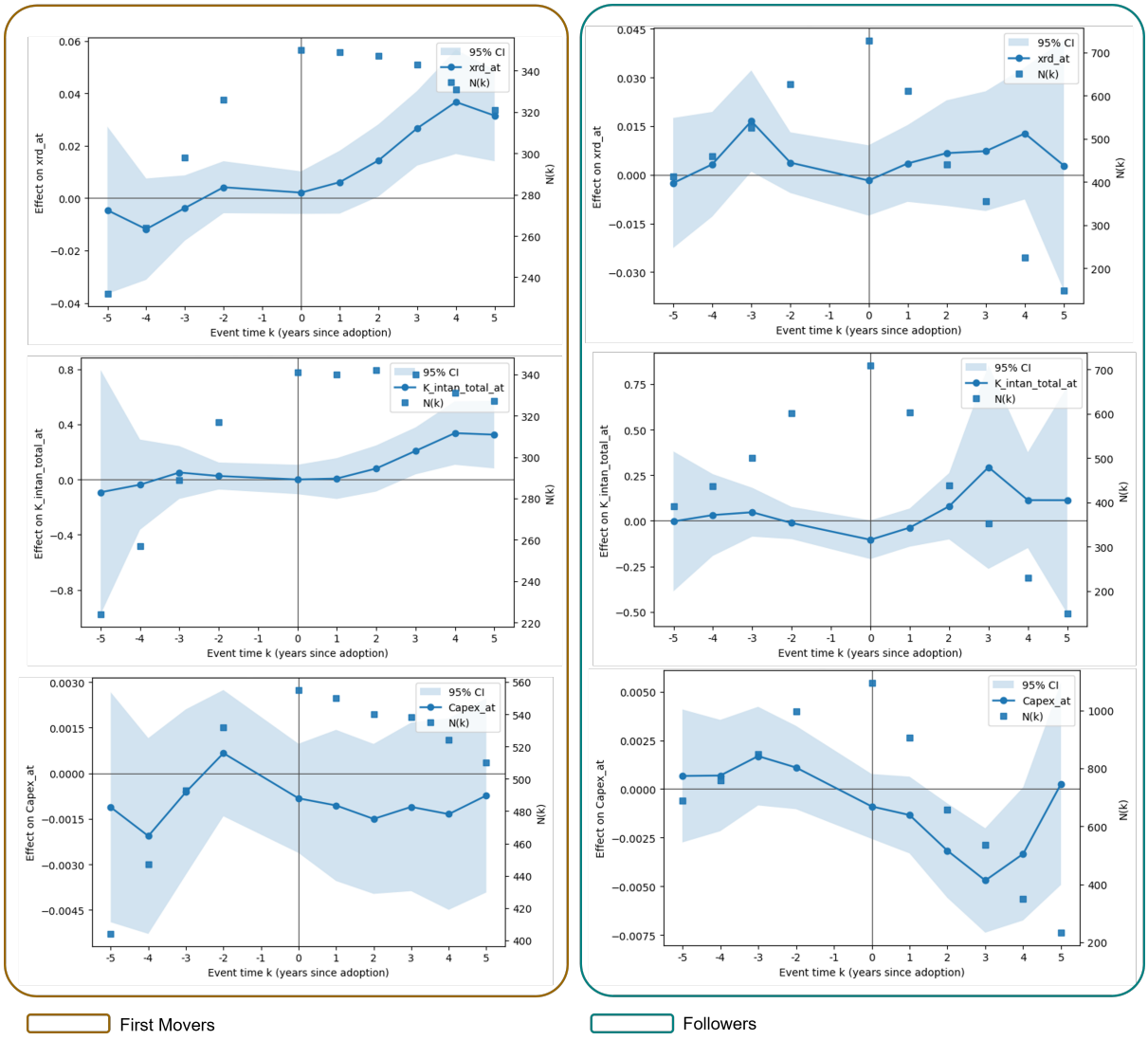


Figure 42: First movers vs. followers: investment responses (Sun–Abraham, IW). Each panel plots Sun–Abraham event-time coefficients with 95% confidence bands (IW implementation), separately for first movers and followers. The top subplot is R&D/Assets and the bottom subplot is Capex/Assets. Event time $k = 0$ denotes the adoption year. $N(k)$ reports the number of observations contributing to each event time.

The predetermined split yields the same mechanism-consistent ranking as the calendar-vintage split: first movers exhibit a substantially stronger post-adoption increase in R&D intensity, whereas followers show weaker R&D build-out and a deeper post-adoption Capex decline (around $k \approx 2$–4). This pattern is consistent with earlier adopters choosing higher installation/build-out effort I_{it} that accumulates into organizational/intangible capital K^I_{it}, while later adopters are closer to "commodity" or defensive adoption where incremental capability accumulation is smaller.

L.2 Diffusion × rents: long-run valuation revaluation

We implement a reduced-form cross-sectional check that directly quantifies rent compression with diffusion. For each adopting firm, define a long-run change in valuation as

$$\Delta q_i = \overline{q}_{i,\,k\in\{3,4,5\}} - \overline{q}_{i,\,k\in\{-2,-1\}},$$

and estimate

$$\Delta q_i = \beta_1 \,\text{Diffusion}_i + \beta_2 \,\text{RentProxy}_{i,-1} + \beta_3 \,\text{Diffusion}_i \times \text{RentProxy}_{i,-1} + \Gamma' X_{i,-1} + \text{Industry FE} + \text{Cohort FE} + \varepsilon_i$$

where $\text{Diffusion}_i \equiv S^{-i}_{g(i),\,G_i-1}$ is the leave-one-out industry adoption share measured pre-adoption, rent proxies are measured pre-adoption (at $k=-1$ when available), and $X_{i,-1}$ includes pre-adoption controls. Industry fixed effects are defined at the adopter's industry, and cohort fixed effects are adoption-year fixed effects. Standard errors are clustered by adoption-industry. Coefficients are interpreted as descriptive associations rather than causal effects.

Using broad intangible capital intensity and R&D intensity as rent proxies, the interaction term is large, negative, and statistically strong across both valuation outcomes: firms with higher pre-adoption "capability" experience smaller long-run revaluation when adoption occurs in high-diffusion industries, consistent with rent compression as diffusion erodes the incremental value of adoption. In contrast, the book-intangibles proxy produces a positive interaction, reinforcing the measurement point emphasized in the main text: accounting intangibles are an imperfect proxy for internally created organizational capital and may reflect acquisition-driven intangibles with different economic content. Accordingly, we treat $K^{\text{intan}}/\text{Assets}$ and $R\&D/\text{Assets}$ as the more mechanism-relevant objects for capability-driven quasi-rents.

M Industry splits

This appendix reports Sun–Abraham event-study estimates separately for the Information Technology and Financials sectors. To facilitate direct comparisons of magnitudes, we

Table 19: Diffusion × rents predicts weaker long-run valuation gains

	Rent proxy (pre-adoption)		
	K^{intan}/Assets	$R\&D$/Assets	Book intangibles
Panel A: Outcome Δq_alt_ev (long-run 3–5 minus pre -2– -1)			
Diffusion	4.6716** (2.0240)	2.8179 (3.4027)	-4.4053* (2.4327)
Rent proxy	0.8308* (0.4620)	-0.0426 (3.3468)	-0.0576 (0.8019)
Diffusion × Rent proxy	-6.4098*** (1.4744)	-40.4766** (18.5433)	9.2373*** (3.3903)
Pre log(mktcap)	-0.0850* (0.0469)	-0.0974** (0.0389)	-0.0802** (0.0317)
Pre operating margin	1.9242** (0.9640)	1.2307 (1.0506)	1.9556 (1.2208)
N firms	548	564	900
Panel B: Outcome Δ Tobin's q (long-run 3–5 minus pre -2– -1)			
Diffusion	2.3991 (1.7989)	2.3341 (2.0940)	-3.2789* (1.9364)
Rent proxy	0.3892 (0.2633)	0.5030 (1.6651)	-0.1238 (0.5626)
Diffusion × Rent proxy	-3.4935*** (0.8068)	-25.4587*** (9.6486)	8.1773*** (2.8075)
Pre log(mktcap)	-0.0839** (0.0423)	-0.0754* (0.0415)	-0.0709** (0.0304)
Pre operating margin	1.1399* (0.6655)	0.8914 (0.5532)	1.3517** (0.6637)
N firms	548	564	900

Notes. Each column reports a separate cross-sectional regression of firm-level long-run valuation changes on pre-adoption diffusion, a pre-adoption rent proxy, and their interaction, controlling for pre-adoption size and profitability. Industry fixed effects are defined at the adoption industry; cohort fixed effects are adoption-year fixed effects. Standard errors are clustered by adoption-industry. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

impose a common y-axis scale across the two sectors. The sector splits provide suggestive evidence on (i) where AI-related costs show up in financial statements and (ii) how factor shares and intangible accumulation adjust following adoption.

In Information Technology (Figure 43), post-adoption effects are positive for both COGS/Sales and SG&A/Sales, consistent with a combination of product-embedded inference/serving costs and deployment/scaling expenditures. IT also exhibits a pronounced and monotone increase in Intangibles/Assets after adoption, indicating systematic intangible accumulation. R&D/Assets dips immediately around $k = 0$ but recovers over the medium run, suggesting that the main post-adoption adjustment is not a mechanically rising R&D share but rather broader intangible formation and operating expenditure.

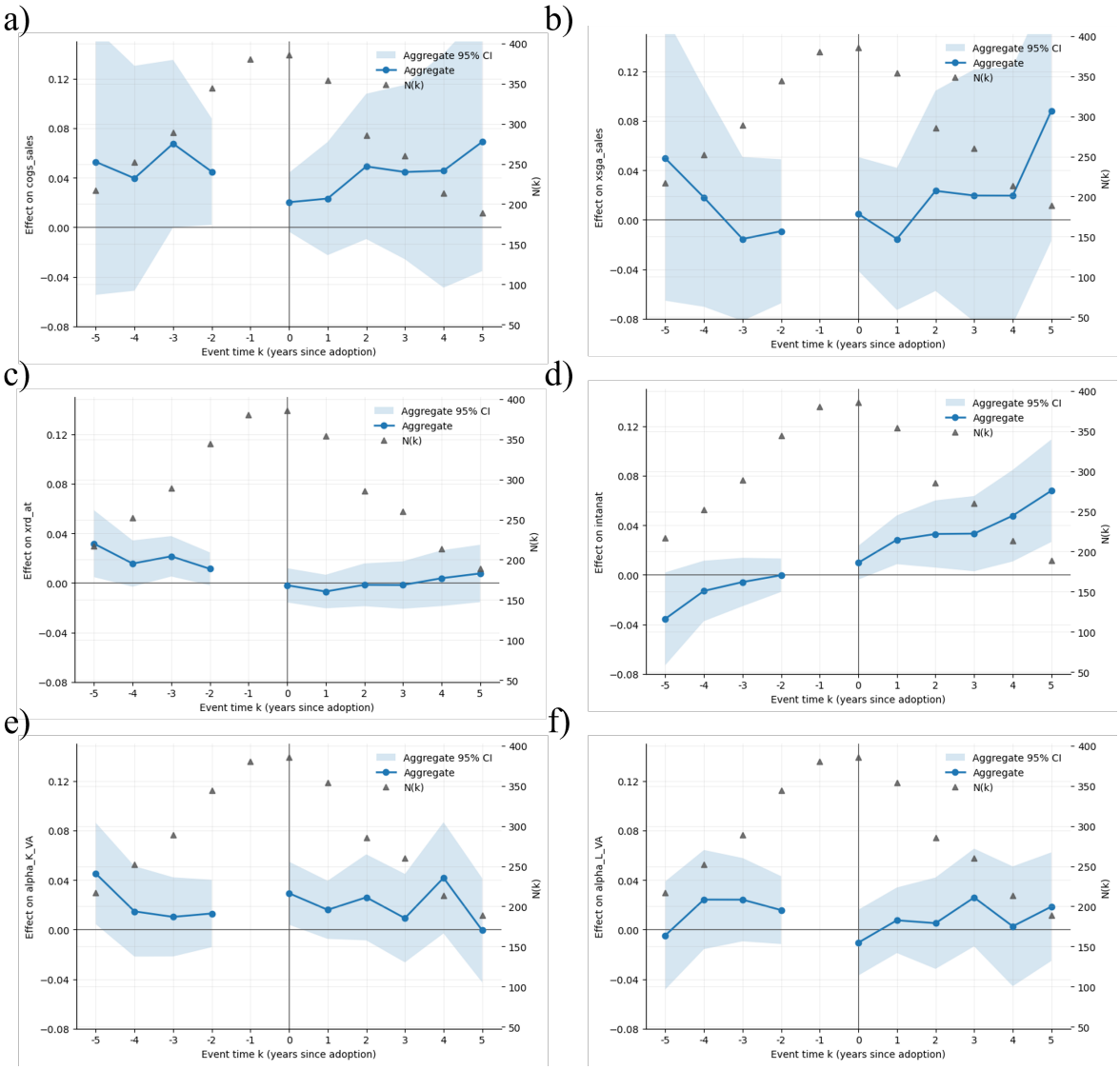


Figure 43: Industry splits (Sun–Abraham): Information Technology. Each panel plots Sun–Abraham event-time coefficients estimated within the IT sector, with 95% confidence bands. The right axis reports the number of observations contributing to each event time, $N(k)$. Outcomes include COGS/Sales and SG&A/Sales (operating cost channels), R&D/Assets and Intangibles/Assets (intangible investment/accumulation), and factor-share measures (capital share and labor share). Event time $k = 0$ denotes the adoption year. Axis limits are standardized across sectors to enable direct magnitude comparisons.

In Financials (Figure 44), the post-adoption pattern differs sharply. SG&A/Sales rises substantially over event time, while COGS/Sales does not increase and is negative in the post period. This pattern is consistent with AI adoption operating primarily through internal automation, reorganization, and compliance/governance costs rather than through product-embedded inference that would load into cost of revenue. Financials also exhibit pronounced factor-share reallocation: the capital share increases markedly post-adoption

while the labor share declines, consistent with automation-driven substitution and/or a shift toward capital-like inputs.

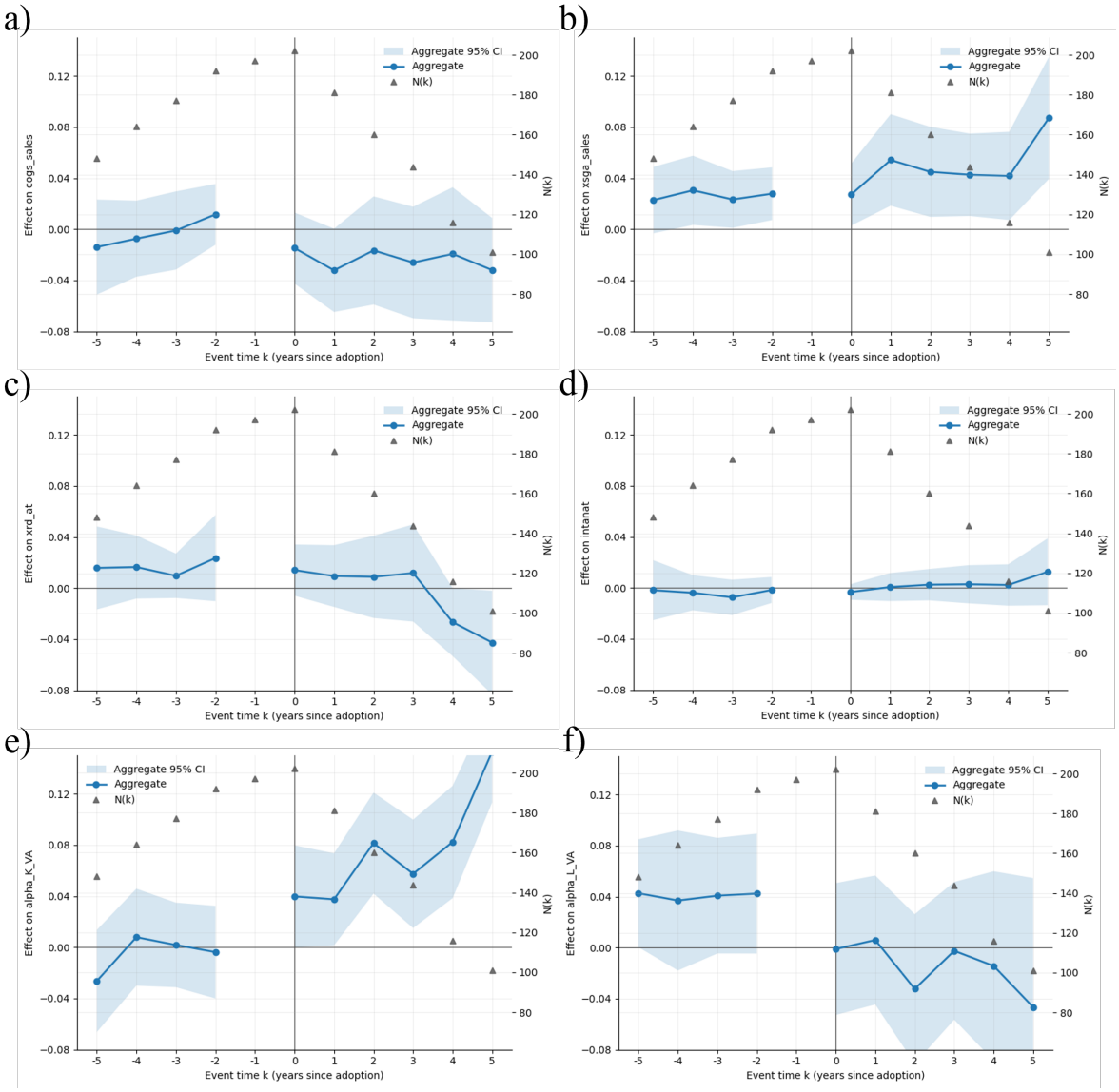


Figure 44: Industry splits (Sun–Abraham): Financials. Each panel plots Sun–Abraham event-time coefficients estimated within the Financials sector, with 95% confidence bands. The right axis reports the number of observations contributing to each event time, $N(k)$. Outcomes include COGS/Sales and SG&A/Sales (operating cost channels), R&D/Assets and Intangibles/Assets (intangible investment/accumulation), and factor-share measures (capital share and labor share). Event time $k = 0$ denotes the adoption year. Axis limits are standardized across sectors to enable direct magnitude comparisons.

N An Interpretation Framework

This section provides a compact interpretation lens—not a structural model and not estimated—that maps directly to four empirical patterns: (i) COGS/Sales rises at adoption while production-based markups do not rise; (ii) SG&A/Sales increases for frontier adopters but decreases for catch-up adopters; (iii) the labor share declines as intangible/organizational capital accumulates; and (iv) the valuation response to adoption declines with industry diffusion.

Adoption and diffusion. Let $D_{it} \in \{0, 1\}$ indicate whether firm i has adopted AI by year t (an absorbing state), with $D_{it} = \mathbf{1}\{t \geq G_i\}$ where G_i is the adoption year. Let $S_{g,t} \in [0, 1]$ denote the diffusion state in industry g, interpreted as the share of firms in g that have adopted by t. In the empirics we operationalize diffusion using the lagged leave-one-out analog

$$S^{-i}_{g(i),t-1} = \frac{1}{N_{g(i),t-1} - 1} \sum_{\substack{j \in g(i) \\ j \neq i}} D_{j,t-1}, \tag{N.1}$$

where $N_{g,t}$ is the number of firms in industry g observed in year t. This is a measurement object; the framework does not by itself provide an exclusion restriction.

Cost shares with non-rising markups. Let R_{it} denote revenue. Let μ_g summarize product-market power and interpret μ_g as the production-based markup measure used in our estimates. Consistent with the evidence, we impose that markups do not rise at adoption:

$$\Delta\mu_{it}\big|_{t=G_i} \leq 0. \tag{N.2}$$

Cost shares can nevertheless move because accounting cost categories (COGS vs. SG&A) respond to adoption through usage costs, expensed build-out, overhead dilution, and the gradual installation of intangibles. A parsimonious decomposition that maps to the data is:

$$\frac{\text{COGS}_{it}}{R_{it}} = \frac{1}{\mu_{g(i)}} + \psi D_{it} + \frac{F^C_i}{R_{it}} - \chi_C \frac{K^I_{it}}{R_{it}}, \qquad \frac{\text{SG\&A}_{it}}{R_{it}} = \phi_{g(i)} + \frac{F^G_i}{R_{it}} - \chi_G \frac{K^I_{it}}{R_{it}} + \frac{I_{it}}{R_{it}}, \tag{N.3}$$

where $\psi > 0$ captures recurring AI usage costs (e.g., purchased model-serving/inference, indicator-driven vendor fees, data/monitoring costs) recorded in COGS and scaling with activity; I_{it} is expensed build-out (typically in SG&A); F_i^C, F_i^G are quasi-fixed components diluted by scale; and K_{it}^I is installed intangible/organizational capital that lowers unit costs broadly.[6]

Frontier vs. catch-up and the SG&A sign. To capture the empirical pattern that SG&A/Sales rises for frontier adopters but falls for catch-up adopters, we impose that expensed build-out intensity is higher when diffusion is low during an installation window:

$$\frac{I_{it}}{R_{it}} = \iota_0 + \iota_1\big(1 - S_{g,t}\big) \cdot \mathbf{1}\{t \in [G_i, G_i + \tau]\}, \qquad \iota_1 > 0. \tag{N.4}$$

When $S_{g,t}$ is low (frontier), the build-out spike is larger and can dominate overhead dilution and efficiency gains, raising SG&A/Sales. When $S_{g,t}$ is high (catch-up), the spike is smaller so dilution and efficiency dominate, lowering SG&A/Sales.

Intangibles and factor shares. Installed intangible capital accumulates according to

$$K_{i,t+1}^I = (1 - \delta_I)K_{it}^I + I_{it}, \tag{N.5}$$

and the key reduced-form implication is that the value-added labor share declines with K^I/L:

$$\alpha_{L,it}^{VA} \equiv \frac{w_t L_{it}}{VA_{it}} \text{ is decreasing in } \frac{K_{it}^I}{L_{it}}. \tag{N.6}$$

Valuation gradient with diffusion. Let q_{it} denote Tobin's q. The framework rationalizes the declining valuation response to adoption with a single monotonicity: the incremental rents (or growth-option value) associated with adopting AI are decreasing in industry diffusion $S_{g,t}$, reflecting competitive catch-up and rent compression. Because installed intangibles accumulate gradually via (N.5), valuation effects can also be delayed as K^I builds post-adoption.

[6] Because μ_{it} is a production-based markup estimator, movements in accounting gross margins induced by ψD_{it} need not imply rising estimated markups; the mechanisms operate through cost composition and accounting classification rather than increased product-market power.

This framework implies four sign restrictions that guide the empirical interpretation:

1. **COGS/Sales:** $\Delta(\text{COGS}/R) > 0$ at adoption through $+\psi D_{it}$ even if (N.2) holds.
2. **SG&A/Sales:** $\Delta(\text{SG\&A}/R)$ is more likely positive when $S_{g,t}$ is low (frontier) and more likely negative when $S_{g,t}$ is high (catch-up) via (N.4).
3. **Factor shares:** labor share declines post-adoption as K^I/L rises through (N.5)–(N.6).
4. **Valuation:** the adoption-related revaluation in q is decreasing in $S_{g,t}$, with the possibility of delayed gains as K^I accumulates.