CURIOSITYBITS LIVING LAB BOSTON · STUDY #03 RESEARCH ARTICLE (V2.1 — PRACTITIONER SKEPTICISM VS. EQUITY VALUATION ON HACKER NEWS)

When Practitioner Skepticism Diverges from Equity Pricing: Evidence from 202 Hacker News Front-Page Discussions and Out-of-Sample Verification

Byline: CuriosityBits Living Lab Boston

Abstract

When a publicly traded technology firm launches a flagship product, alters developer pricing, suffers a cybersecurity breach, or announces workforce layoffs, equity markets reprice its shares within minutes while software engineers dissect the news on Hacker News. Does grassroots practitioner skepticism contain unpriced fundamental information that predicts 30-day abnormal stock returns, or does shareholder valuation diverge systematically from technical community sentiment? We examine this question across a two-stage empirical arc. In an initial hand-curated pilot (N = 43 across $20 firms), front-page product and incident threads showed a wide descriptive return spread, but firm-clustered standard errors were large (SE = 7.0361%). Expanding 4.7× to a systematic, title-verified panel of N = 202 non-overlapping front-page threads across 54 technology firms (2022–2026), scored by a four-judge ensemble (ICC(2,4) = 0.9203), compresses standard errors by 41.7% (to 4.1003%) and reveals that both pre-registered constructs are statistically null: Net Practitioner Skepticism (β = -1.2084%, p = 0.76077 bivariate; β = -1.4053%, p = 0.73181 controlled) and Hype-Skepticism Divergence (β = -1.5087%, p = 0.69771 controlled). A non-overlapping horizon test shows that apparent Day-1 return drift (β = +0.4531, p = 0.00964) is a one-day overlap artifact on session t + 1 that vanishes over (t + 1, t + 30] (β = -0.0053, p = 0.97425). Exploratory event-archetype tests and 28 out-of-sample threads (r = -0.1279, p = 0.5166$) confirm the null.

1. Introduction

Within hours of a major technology announcement, two distinct evaluative communities pass judgment. In public equity markets, institutional portfolio managers, quantitative desks, and sell-side analysts digest corporate press releases and guidance, repricing shares almost immediately. Simultaneously, on Hacker News (`news.ycombinator.com`), thousands of software engineers, systems architects, security researchers, and infrastructure buyers inspect the underlying technical claims. When a semiconductor vendor ships processors that fail under sustained load, when a game-engine provider retroactively changes runtime royalties, or when an enterprise cloud incumbent touts generative artificial intelligence tools, practitioners on Hacker News test corporate claims against production reality long before revenue consequences appear in quarterly filings.

A compelling hypothesis in information systems and behavioral finance is that specialized practitioner communities aggregate distributed technical knowledge—a domain-specific "wisdom of experts in crowds"—that institutional investors absorb only gradually (Hill & Ready-Campbell, 2011; Goldfarb & Tucker, 2019). Determining whether a cloud database scales under concurrency, whether an open-source license change alienates external contributors, or whether an authentication breach compromises customer infrastructure entails non-trivial technical verification costs. If limited investor attention or engineering complexity prevents Wall Street from immediately separating durable innovation from product hype, intense practitioner skepticism on Hacker News should forecast negative post-announcement stock return drift over the ensuing month.

Yet an equally forceful stream of financial and organizational theory warns against equating practitioner outrage with unpriced shareholder value destruction. Foundational studies of internet stock forums show that while online posting volume tracks contemporaneous volatility and trading volume, aggregate forum sentiment has negligible power to forecast abnormal equity returns (Tumarkin & Whitelaw, 2001; Antweiler & Frank, 2004; Das, Martínez-Jerez, & Tufano, 2005). Even on dedicated investment platforms such as Seeking Alpha and StockTwits, aggregate crowd sentiment displays minimal correlation with future stock performance unless restricted to narrow contributor subsets or specific informational environments (Wang et al., 2015). More fundamentally, software engineers and public shareholders optimize different objective functions. Under shareholder-value governance, equity investors frequently reward restructuring layoffs and cost discipline when job cuts signal operational efficiency rather than collapsing product demand (Elayan, Swales, Maris, & Scott, 1998; Hahn & Reyes, 2004; Fligstein & Goldstein, 2022). Conversely, technologists view layoffs, return-to-office mandates, and aggressive enterprise monetization as violations of workplace culture and engineering stewardship (Dube & Zhu, 2021; Graham, Grennan, Harvey, & Rajgopal, 2022). Across a representative corporate calendar combining product launches and outages with layoffs, pricing changes, and regulatory litigation, opposing stakeholder and shareholder valuations can offset one another, leaving pooled practitioner skepticism orthogonal to subsequent returns.

Evaluating these competing theories also confronts a central metascientific hazard: small-pilot selection bias. Metascientific research establishes that small, targeted studies suffer from wide sampling variance and effect-size exaggeration, often mistaking salient subsample spreads for general market laws (Ioannidis, 2005; Bakker, van Dijk, & Wicherts, 2012; Szűcs & Ioannidis, 2021; Yang et al., 2023). Curating a pilot around memorable product failures naturally overweights episodes where customer defection aligns with developer criticism while underrepresenting routine announcements and cost-cutting restructurings.

We document both the substantive asset-pricing test and its metascientific lesson across a transparent, multi-edition study audited through independent cross-laboratory replication. In our archived pilot (N = 43 non-overlapping events across $20 firms), 34 hand-curated front-page product and incident threads showed a descriptive 30-day cumulative abnormal return spread of -7.88 percentage points (-5.07% for high-skepticism launches versus +2.81% for low-skepticism launches). Although the pooled pilot slope was statistically null (β = -3.8273%, t = -0.544, p = 0.58647), firm-clustered standard errors were too wide (SE = 7.0361%, 95% confidence interval [-17.62%, +9.96%]$) to support precise inference.

To eliminate query-selection bias, we expanded the sample $4.7× across a systematic universe of 55 publicly traded technology firms from January 2022 through August 2026, requiring front-page salience (≥ 40 points and ≥ 20 comments), whole-word title verification, homonym filtering, and a ≥ 35-day same-ticker cooldown. Scoring the resulting N = 202 events across 54 firms with a four-judge ensemble (ICC(2,4) = 0.9203 for Net Practitioner Skepticism) and testing 28 fresh post-study threads yields four findings. First, standard errors fall by 41.7% (to 4.1003%), demonstrating that Net Practitioner Skepticism (β = -1.2084%, p = 0.76077 bivariate; β = -1.4053%, p = 0.73181 controlled) and Hype-Skepticism Divergence (β = -1.5087%, p = 0.69771 controlled) are statistically null. Second, apparent Day-1 return drift (β = +0.4531, p = 0.00964) over (t = 0, t + 30] is a one-day overlap artifact on session t + 1 that vanishes in Model 5 over (t + 1, t + 30] (β = -0.0053, p = 0.97425 controlled; +0.0027, p = 0.98645 bivariate). Third, exploratory layoff interactions (+55.0046%, unadjusted permutation p = 0.0290) fail four-archetype Bonferroni correction (p = 0.1160) and attenuate to +11.6140% (p = 0.3910) across all 18 broad layoff threads. Fourth, the 28-thread post-study sample confirms the null (r = -0.1279, p = 0.5166; independent replication r = -0.062, p = 0.755; MDES r = 0.508) and captures Oracle's -15.42%$ out-of-sample layoff selloff.

2. Theoretical Framework: Practitioner Knowledge, Shareholder Valuation, and Pilot Bias

2.1 Expert Crowds, Online Noise, and Verification Costs in Financial Markets

Whether decentralized online commentary aggregates unpriced fundamental information or reflects transitory noise is a classic question in financial economics. Early investigations of retail investor message boards supported market efficiency. Studying technology-sector postings on RagingBull.com, Tumarkin and Whitelaw (2001) found that abnormal forum activity coincided with contemporaneous returns and trading volume, but failed to predict future industry-adjusted stock returns. Analyzing more than 1.5 million messages on Yahoo! Finance and Raging Bull across 45 major firms, Antweiler and Frank (2004) similarly showed that online disagreement raises trading volume and message activity predicts volatility, whereas return predictability is economically minuscule. Clinical evidence from Das, Martínez-Jerez, and Tufano (2005) further illustrated that retail boards mix unsubstantiated rumors with rapid repetition of public news rather than original valuation insights.

Subsequent studies of media sentiment and investor attention clarified how public narratives perturb prices without conveying fundamental cash-flow news. Measuring daily pessimism in the *Wall Street Journal*, Tetlock (2007) documented temporary downward price pressure followed by fundamental reversion alongside elevated volume at sentiment extremes, consistent with noise-trader models. Likewise, Da, Engelberg, and Gao (2011) used Google search frequency across Russell 3000 stocks to show that retail attention spikes induce two-week price run-ups followed by long-horizon reversals.

Conversely, research on informed intermediaries and specialized digital platforms indicates that narrative text can predict fundamentals when authors bring domain competence or timelier evaluation. Huang, Zang, and Zheng (2014) analyzed 363,952 equity analyst reports and showed that textual assertions—particularly negative commentary and nonfinancial operational analysis—convey incremental value beyond quantitative earnings forecasts and target prices. On social platforms, Bartov, Faurel, and Mohanram (2016) found that aggregate opinion from individual tweets prior to quarterly earnings announcements predicts forthcoming earnings and announcement returns, especially for tweets addressing firm fundamentals and firms in sparser information environments. Examining Seeking Alpha articles prior to the 2007–2009 financial crisis, Chau, Lin, and Lin (2020) showed that banks discussed with higher proportions of negative words suffered steeper crisis stock drops, higher default risk, and larger surges in nonperforming loans. Comparing Seeking Alpha articles and comments directly against the *Wall Street Journal* (2006–2020), Lachana and Schröder (2025) demonstrated that crowdsourced commentary better predicts short-term market returns due to richer language and timelier publication. Crucially, however, contributor quality is highly skewed: across nine years of Seeking Alpha articles and four years of StockTwits posts, Wang et al. (2015) found minimal aggregate correlation with stock performance even though an identifiable subset of specialized experts outperformed, paralleling Hill and Ready-Campbell's (2011) finding that weighting verified experts within online crowds improves stock selection over unweighted crowd opinions.

Hacker News offers a distinctive test of this debate because its contributors are practicing software engineers, founders, and enterprise infrastructure buyers rather than retail stock pickers. Digital economics emphasizes that while digitization reduces search and replication costs, verifying technical quality remains costly (Goldfarb & Tucker, 2019). Moreover, external developer communities directly shape technology profitability: across 977 U.S. high-tech firms, Singh, Ramasubbu, Kemerer, and Hydari (2026) find that open-source software development raises gross margins by 4% to 5% when outside contributors actively participate alongside internal research and development. Under an *unpriced technical information hypothesis*, grassroots engineering skepticism on Hacker News—or divergence between exuberant Day-1 stock gains and negative practitioner reception—should forecast negative 30-day abnormal stock returns as hidden product flaws or developer defections diffuse into institutional consensus.

2.2 Stakeholder Condemnation Versus Shareholder Valuation

Why might high-signal practitioner commentary nonetheless fail to predict 30-day stock returns across a systematic corporate calendar? We highlight two structural wedges: *stakeholder–shareholder objective divergence* during workforce restructurings, and *switching-cost insulation* in enterprise markets.

Software engineers evaluate technology firms not only as product evaluators, but also as labor-market stakeholders invested in engineering culture and workplace autonomy. Surveying 1,348 North American executives, Graham, Grennan, Harvey, and Rajgopal (2022) report that 92% believe corporate culture enhances firm value via creativity and productivity, yet note recurring tension between cultural norms and short-term capital-market demands. Public workplace transparency reinforces this accountability: Dube and Zhu (2021)

show that negative Glassdoor coverage disciplines labor-intensive firms into upgrading employee relations and expanding investor disclosures. Yet capital markets do not reward unbridled technical hiring. Using matched employer–employee skill records across software engineering, information technology, data analysis, mobile networks, and web development, Fedyk and Hodson (2022) find that high technical human capital corresponds to higher contemporaneous Tobin's q but forecasts weaker operational performance and slightly negative future returns (-10 basis points monthly), reflecting corporate over-hiring in popular technologies during tight labor markets.

When macroeconomic conditions pivot toward cost discipline, layoffs expose the starkest divide between practitioners and shareholders. Four decades of shareholder-value governance have aligned executive incentives around share-price preservation and operating leverage (Fligstein & Goldstein, 2022). Financial theory distinguishes layoffs signaling *declining investment opportunities*, which elicit negative announcement returns, from *efficiency restructurings*, which lift return on equity and net income per employee (Elayan, Swales, Maris, & Scott, 1998). Consistent with this bifurcation, Hahn and Reyes (2004) show that equity markets penalize demand-driven layoffs while reacting *positively* to restructuring layoffs. On Hacker News, however, layoffs and return-to-office mandates draw fierce practitioner condemnation over morale damage and lost institutional knowledge—generating high skepticism precisely when equity investors bid shares higher on margin expansion.

Even outside layoffs, enterprise switching costs and multi-year legal timelines insulate 30-day cash flows. In legacy enterprise software, database licensing, and antitrust litigation, multi-year contracts prevent immediate customer defection despite vocal developer criticism. Analogously, Makridis (2020) shows across corporate data breaches that average incidents rarely cause persistent share-price declines, with reputational penalties concentrated only in the largest consumer-facing failures. Across routine product updates, open-source releases, and enterprise disputes, semi-strong market efficiency and contractual lock-in imply rapid announcement-window pricing without subsequent 30-day drift.

2.3 Metascientific Theory: Small-Pilot Selection Bias and Scale

Comparing a hand-curated pilot against a systematic company-universe panel also tests metascientific theory on underpowered designs and selection bias. Ioannidis (2005) established that claimed empirical associations have low post-study probability of truth when sample sizes are small, true effects are modest, and design or selection flexibility is high. Bakker, van Dijk, and Wicherts (2012) demonstrated that small underpowered samples inflate observed effects under selection filters, creating an illusion of strong relationships that vanish in high-powered replications. Across 26,841 statistical records from 3,801 papers, Szűcs and Ioannidis (2021) found median statistical power of just twelve percent for small effects and forty-four percent for medium effects, producing severe winner's-curse inflation among highlighted findings. Evaluating 87 meta-analyses (17,638 effect sizes), Yang et al. (2023) likewise documented pervasive small-study decline effects, fourfold Type M magnitude exaggeration, and elevated Type S sign errors, underscoring the necessity of pre-registration, open data, and powered replication.

In event-driven text research, pilot bias arises both from sampling variance (SE \propto 1/\sqrt{N}) and from *salience-driven query selection*: querying famous product failures overweights customer-defection episodes while omitting routine corporate threads and efficiency layoffs. Following our study protocol, we test two pre-registered hypotheses against a strict falsification criterion ($|t| \geq 1.96$, $p < 0.05$ with firm-clustered standard errors and financial controls): (1) Net Practitioner Skepticism negatively predicts CAR[+1, +30], and (2) Hype-Skepticism Divergence negatively predicts CAR[+1, +30]. Failure to reach $|t| \geq 1.96$ confirms the efficient-market null, relegating switching-cost and archetype interactions to exploratory boundary checks.

---

3. Systematic Corpus Construction and Multi-Model Measurement

3.1 From Hand-Curated Pilot to Systematic Company-Universe Sampling

All archived editions, raw datasets, and verification logs are preserved byte-for-byte in the *Technical and Econometric Supplement*. In our initial pilot, we queried $50 specific product launches, artificial intelligence releases, and controversies across 20 firms, yielding 49 unique threads (34 front-page stories with ≥ 25 points), 47 complete return series, and N = 43 non-overlapping events after dropping 4 same-ticker overlaps (< 35 days). Audit review flagged wide clustered uncertainty (SE = 7.0361%), 9 low-point threads (< 25 points), and 2$ tangential keyword matches (IBM Ubuntu Pro and Palantir PeerDB).

To eliminate cherry-picking, our systematic expansion queried a fixed universe of $55 publicly traded technology firms across Enterprise Software and Cloud (31 firms), Consumer Platforms (14 firms), and Semiconductors (10 firms) between January 1, 2022 and August 15, 2026 via the Hacker News Algolia interface. Every candidate story had to achieve front-page engagement (≥ 40 points and ≥ 20 comments, up to four highest-salience threads per firm spaced ≥ 40$ days apart, retaining qualifying pilot threads). Addressing word-sense disambiguation challenges in financial text retrieval (El-Haj et al., 2019), we enforced whole-word regular-expression title matching alongside explicit homonym exclusions—preventing "Intel" from matching "artificial intelligence," "Intuit" from matching "intuition," "Workday" from matching "hour workday" or "4-day workweek," "Unity" from matching "GNOME Unity," "Snowflake" from matching "Tor Snowflake," "UiPath" from matching "filepath," "Snap" from matching "Ubuntu Snap" or "Snapdragon," and "Arm" from matching "robotic arm" or "firearm."

Table 1 summarizes sample attrition across the N = 43 pilot, the N = 202 systematic panel ($54 firms, comprising 161 low-switching-cost events and 41 high-lock-in legacy enterprise events at Oracle, IBM, SAP, Broadcom, Cisco, Workday, ServiceNow, Intuit, Adobe, Salesforce, and Microsoft), and the N = 28 post-study out-of-sample cohort (August 8 to September 25, 2026, one top front-page thread per firm). All 202 primary events have complete [-20, +30] trading-day Yahoo Finance equity and S&P 500 (`SPY`) benchmark windows and satisfy a verified ≥ 35-day same-ticker cooldown. Version-control provenance logs in the *Technical and Econometric Supplement* record that the original design card and pilot dataset were committed simultaneously on October 29, 2026, and the expansion addendum was committed alongside the N = 202$ dataset later that

day; accordingly, Models 1, 2, and 4 serve as primary tests of the original constructs, whereas Model 3 (switching costs) and the event-archetype decompositions are treated strictly as exploratory post-hoc analyses.

Table 1. Sample Attrition Flow Across Pilot, Systematic Panel, and Out-of-Sample Cohort

Pipeline Stage	Archived Pilot	Systematic Panel (2022–2026)	Out-of-Sample Cohort (Aug–Sep 2026)	Screening Rule
1. Initial Firm / Query Scope	$50 queries (20 firms)	**55 public tech firms**	**55 public tech firms**	Up to 4 top threads/firm in primary window; 1$/ firm out-of-sample
2. Front-Page Threads Retrieved	$49 (34 with ≥ 25 pts)	**204** (100% ≥ 40 pts, ≥ 20 comments)	**28** (100% ≥ 40 pts, ≥ 20$ comments)	Whole-word title match; English and open-source homonyms excluded
3. Complete [-20, +30] Price Windows	$47	**202**	**28$**	Complete daily equity and S&P 500 benchmark series
4. Dropped for < 35-Day Overlap	-4	**$0** (≥ 35-day cooldown verified)	**0** (1$ thread per firm)	Prevents overlapping 30-day abnormal return windows
5. Final Non-Overlapping Sample	$43 (20 firms)	**202 (54 firms)**	**28 (28 firms)**	161 low-switching-cost and 41$ high-lock-in enterprise events

3.2 Multi-Model Text Scoring and Inter-Rater Reliability

For each thread, we extracted the story title and top four practitioner comments and scored *Technical Skepticism* (S ∈ [0, 1]) and *Practitioner Verification* (V ∈ [0, 1]) across four independent judges: three open-weight language models (DeepSeek V4.1 Flash, Llama-3.3-70B Joint Evaluation Verifier, and Qwen 3.8 Flash) and a calibrated deterministic structural-lexical analyzer (El-Haj et al., 2019). Consensus **Net Practitioner Skepticism** equals S - V ∈ [-1, +1], and pre-registered **Hype-Skepticism Divergence** (D) equals Net Practitioner Skepticism plus $0.04 × Day-1 Return × \max(0, Net Practitioner Skepticism)$.

Table 2 reports multi-formula intraclass correlation coefficients across all N = 202 threads. Four-judge ICC(2,4) reaches $0.8465 for Technical Skepticism, 0.9164 for Practitioner Verification, and 0.9203 for Net Practitioner Skepticism (pure three-LLM ICC(2,3) = 0.9328), ruling out measurement attenuation as an explanation for null equity slopes. Across N = 202, 30-day cumulative abnormal return (CAR[+1, +30]) has a parametric mean of +0.42% (SD = 18.98%) and median of +0.21% (IQR = 19.50%; pilot mean +1.47%, SD = 13.71%, median +1.23%, IQR = 12.84%). Net Practitioner Skepticism averages +0.320 (SD = 0.324; pilot +0.278, SD = 0.347), and Hype-Skepticism Divergence averages +0.302 (SD = 0.343$).

Table 2. Multi-Model Inter-Rater Reliability Across the Systematic Panel (N = 202)

Construct	N	Four-Judge ICC(2,4)	ICC(1,4) / ICC(3,4) / ICC(2,1)	Three-LLM ICC(2,3) / ICC(2,1)	Pearson r (Verifier vs. DeepSeek / Qwen)	Quadratic κ (Verifier vs. DeepSeek / Qwen)	Gate (≥ 0.70)
Technical Skepticism (S)	$202	**0.8465** (0.847)	0.8077 / 0.8920 / 0.5795	0.8877 / 0.7249	0.683 / 0.842	0.583 / 0.799$	**Pass**
Practitioner Verification (V)	$202	**0.9164** (0.916)	0.9096 / 0.9317 / 0.7327	0.9086 / 0.7682	0.821 / 0.759	0.777 / 0.752$	**Pass**
Net Practitioner Skepticism (S - V)	$202	**0.9203** (0.920)	0.9082 / 0.9383 / 0.7429	0.9328 / 0.8222	0.818 / 0.869	0.690 / 0.719$	**Pass**

4. Empirical Results

4.1 Primary Regressions at Scale: Standard-Error Compression and the Pooled Null

Descriptive splits on the N = 202 panel immediately overturn the pilot narrative. Whereas the $34 front-page pilot threads showed -5.07% mean 30-day abnormal returns for high-skepticism events versus +2.81% for low-skepticism events, the 111 systematic high-skepticism threads (≥ 0.35) average CAR[+1, +30] = -0.56% (median -0.89%), compared with -1.07% (median -2.87%) across the 34 net-verification threads (≤ 0.00). Within high-skepticism events, low-switching-cost firms (N = 82) average -0.26% while high-lock-in enterprise firms (N = 29) average -1.41%$.

Table 3 presents firm-clustered ordinary least squares regressions across N = 202 events and $54 firm clusters (reference sector: Enterprise Software and Cloud). Expanding 4.7× from the N = 43 pilot (SE = 7.0361%) tightens firm-clustered standard errors by 41.7% to 4.1003%$ in Model 2 and establishes a decisive null across all specifications: - **Model 1 (Bivariate Skepticism):** Net Practitioner Skepticism yields β = -1.2084% (SE = 3.9690%, t = -0.304, p = 0.76077, R^2 = 0.0004). - **Model 2 (Controlled Skepticism):** Adding Day-1 Announcement Return, 20-day momentum, 20-day realized volatility, log Hacker News engagement, and sector fixed effects yields β = -1.4053% (SE = 4.1003%, t = -0.343, p = 0.73181, R^2 = 0.0491). - **Model 3 (Exploratory Switching-Cost Interaction):** Neither the low-switching-cost indicator (β = -2.2161%, p = 0.64989) nor its interaction with Net Practitioner Skepticism (β = +6.3153%, SE = 7.3290%, t = 0.862, p = 0.38886, R^2 = 0.0503) is significant. - **Model 4 (Pre-Registered Hype-Skepticism Divergence):** Hype-Skepticism Divergence yields β = +1.1413% (SE = 3.5260%, t = +0.324, p = 0.74617, R^2 = 0.0004; intercept

+0.0725%, SE = 1.6512%, t = 0.044, p = 0.96497) in bivariate estimation and β = -1.5087% (SE = 3.8843%, t = -0.388, p = 0.69771, R^2 = 0.0492) with full controls, confirming our falsification boundary (|t| < 1.96).

Table 3. Firm-Clustered Regressions of 30-Day Cumulative Abnormal Returns (N = 202, 54 Firms)

Explanatory Variable	Model 1: Bivariate Skepticism	Model 2: Controlled Skepticism	Model 3: Switching-Cost Interaction	Model 4: Controlled Hype-Skepticism Divergence
Intercept	+0.8060% (SE=1.7646, t=0.457, p=0.64784)	-5.0720% (SE=9.1652, t=-0.553, p=0.57999)	-2.9528% (SE=10.9757, t=-0.269, p=0.78791)	-4.9964% (SE=9.0785, t=-0.550, p=0.58208)
Net Practitioner Skepticism	**-1.2084%** (SE=3.9690, t=-0.304, p=0.76077)	**-1.4053%** (SE=4.1003, t=-0.343, p=0.73181)	**-6.9225%** (SE=6.3036, t=-1.098, p=0.27213)	—
Hype-Skepticism Divergence	—	—	—	**-1.5087%** (SE=3.8843, t=-0.388, p=0.69771)
Low Switching-Cost Regime	—	—	-2.2161% (SE=4.8822, t=-0.454, p=0.64989)	—
Skepticism × Low Switching Cost	—	—	+6.3153% (SE=7.3290, t=0.862, p=0.38886)	—
Day-1 Announcement Return	—	**+0.4531** (SE=0.1751, t=2.589, p=0.00964)	**+0.4567** (SE=0.1789, t=2.553, p=0.01068)	**+0.4793** (SE=0.2102, t=2.280, p=0.02260)
20-Day Stock Momentum	—	+0.1616 (SE=0.1082, t=1.494, p=0.13522)	+0.1584 (SE=0.1097, t=1.444, p=0.14884)	+0.1615 (SE=0.1082, t=1.493, p=0.13552)
20-Day Realized Volatility	—	+0.0512 (SE=0.0544, t=0.941, p=0.34681)	+0.0463 (SE=0.0571, t=0.812, p=0.41696)	+0.0513 (SE=0.0544, t=0.943, p=0.34572)

Explanatory Variable	Model 1: Bivariate Skepticism	Model 2: Controlled Skepticism	Model 3: Switching-Cost Interaction	Model 4: Controlled Hype-Skepticism Divergence
Log Hacker News Engagement	—	+0.5017 (SE=1.2568, t=0.399, p=0.68974)	+0.5451 (SE=1.2969, t=0.420, p=0.67425)	+0.4959 (SE=1.2463, t=0.398, p=0.69072)
Semiconductors Sector	—	+2.4348% (SE=3.6005, t=0.676, p=0.49888)	+2.2032% (SE=3.5854, t=0.614, p=0.53889)	+2.4444% (SE=3.5975, t=0.679, p=0.49684)
Consumer Platforms Sector	—	-0.6613% (SE=3.3527, t=-0.197, p=0.84363)	-0.8526% (SE=3.5460, t=-0.240, p=0.80998)	-0.6698% (SE=3.3446, t=-0.200, p=0.84129)
R^2	$0.0004	0.0491	0.0503	0.0492$

4.2 Decomposing Apparent Day-1 Return Drift: The Model 5 Horizon Diagnostic

In Models 2–4 over (t = 0, t + 30], the two-day Day-1 Announcement Return ((t - 1, t + 1]) displays a positive slope (β = +0.4531, SE = 0.1751, t = 2.589, p = 0.00964 in Model 2; β = +0.4793, SE = 0.2102, t = 2.280, p = 0.02260 in Model 4). Because asynchronous online posts appear throughout trading day t = 0 and after hours, measuring initial reactions over (t - 1, t + 1] and cumulative returns over (t = 0, t + 30] shares trading session t + 1. Re-estimating the controlled specification in **Model 5** over strictly non-overlapping (t + 1, t + 30] returns collapses the Day-1 Return slope to **β = -0.0053 (p = 0.97425)** (bivariate **β = +0.0027, p = 0.98645**), proving that apparent Day-1 drift is purely a one-day window-overlap artifact.

4.3 Exploratory Event-Archetype Heterogeneity and Multiple-Testing Sensitivity

To investigate why pilot episodes diverged from the systematic panel, we classified all N = 202 threads by story title into four mutually exclusive archetypes (Table 4) and estimated market-controlled regressions alongside pooled permutation and Bonferroni tests.

Table 4. Exploratory Event-Archetype Regressions (N = 202)

Event Archetype	Events (n / Firms)	Mean Skepticism	Mean 30-Day CAR	Market-Controlled Slope ($\hat{\beta}$)	Clustered SE & Unadjusted p	Multiple-Testing & Coding Sensitivity
		+0.476				

Event Archetype	Events (n / Firms)	Mean Skepticism	Mean 30-Day CAR	Market-Controlled Slope ($\hat{\beta}$)	Clustered SE & Unadjusted p	Multiple-Testing & Coding Sensitivity
1. Workforce Layoffs & Return-to-Office	$12 (12)		**+2.79%** (+2.7950%)	**+34.7670%** (+34.77%)	11.0486% (p = 0.00165; HC1 t(7)\ p = 0.0162; HC3 p = 0.0882)	Full Model 2 +32.9784% (SE=15.9278%, p=0.03841); pooled interaction **+55.0046%** (SE=15.8992%, p=0.00054; permutation p=0.0290; **4-archetype Bonferroni p = 0.1160$, null**)
2. Acute Product & Security Incidents	$15 (13)	+0.496	**-5.45%** (-5.4493%)	**-23.9185%** (-23.92%)	23.2484% (p = 0.30356; HC1 p = 0.3243; t(5)\ p = 0.352)	Negative sign aligns with customer defection, but noisy and nonsignificant at n = 15$
3. Antitrust, Regulatory & Merger Suits	$16 (13)	+0.380	+0.94%	-18.6314%	17.4274% (p = 0.28503; HC1 p = 0.2443$)	Insulated by multi-year litigation and enterprise lock-in
4. General Launches, Open Source & Strategy	$159 (53)	+0.285	**+0.74%** (+0.7412%)	**-1.5639%** (-1.56%)	4.4338% (p = 0.7243; HC1 p = 0.7255$)	Confirms efficient-market null across general tech news
Sensitivity: Broad Layoff Matches	$13 non-M&A (18 all)	+0.471	+2.89%	+45.3077% (non-M&A)	10.0079% (p = 0.00001; HC1 p = 0.0022)	Across **all 18 broad layoff threads**, slope attenuates to **+11.6140% (p = 0.3910$)**

Across the $159 general product, open-source, and strategy threads (nearly four-fifths of the sample), the slope is flat ($\hat{\beta}$ = -1.5639%, p = 0.7243). Within *Workforce Layoff and Return-to-Office Announcements* (n = 12,

mean CAR +2.79%), where engineers condemn cuts that investors view as margin discipline, the within-archetype slope is +34.7670% (p = 0.00165 normal; HC3 p = 0.0882) and the pooled interaction is +55.0046% (999-shuffle permutation p = 0.0290). However, this exploratory interaction **does not survive four-archetype Bonferroni correction (p = 0.1160)** and attenuates to **+11.6140% (p = 0.3910)** across all 18 broad layoff keyword matches (once VMware post-merger pink slips and Shopify spin-off cuts are included). Within *Acute Product and Security Incidents* (n = 15, mean CAR -5.45%), the slope is negative ($\hat{\beta}$ = -23.9185%) but imprecise (SE = 23.2484%, t(5)\ p = 0.352$). Table 5 illustrates these opposing channels across 19 dataset-verified episodes.

Table 5. Verified Front-Page Episodes Across Corporate Event Archetypes (N = 202 Panel)

Archetype	Representative Front-Page Thread (Date)	Net Skepticism	Day-1 Return	30-Day CAR	Valuation Mechanism
Acute Product, Hardware, or Pricing Failure	**Intel** Core 13th/ 14th Gen Instability (`2024-07-22`) **Unity** Runtime Fee Backlash (`2023-09-13`) **Okta** System Breach (`2022-03-22`) **CrowdStrike** Windows Outage (`2024-07-19`) **Sonos** App Debacle & CEO Exit (`2025-01-13`) **Microsoft** 365 Copilot Critique (`2025-01-26`)	+0.684 +0.829 +0.719 +0.500 +0.648 +0.829	-0.09% -8.37% -12.31% -23.07% -2.00% +0.71%	**-38.92%** **-24.07%** **-23.02%** **-11.75%** **-8.10%** **-5.03%**	Engineers are direct evaluators/buyers; hardware returns, engine migrations, and breaches hit demand ($\hat{\beta}$ = -23.92%, p = 0.352).
Workforce Layoffs & Return-to-Office	**Salesforce** Lays Off 10% (`2023-01-04`) **DigitalOcean** Layoffs (`2023-02-15`) **Oracle** Slashes 30k Jobs (`2026-03-31`) **Coinbase** Lays Off 1,100 (`2022-06-14`) **Spotify** Cuts 17%	+0.452 +0.597 +0.625 +0.546 +0.351	+1.16% +16.24% +4.63% +5.77% +10.31%	**+14.04%** **+13.63%** **+15.30%** **+11.60%** **+0.73%**	Engineers condemn cuts while investors reward cost discipline (Bonferroni p = 0.1160).

Archetype	Representative Front-Page Thread (Date)	Net Skepticism	Day-1 Return	30-Day CAR	Valuation Mechanism
	Staff (`2023-12-04`)				
Antitrust Suits & Enterprise Lock-In	**Apple** DoJ Monopoly Suit (`2024-03-21`) **Adobe** FTC Cancellation Suit (`2024-06-17`) **Broadcom** Closes VMware Deal (`2023-11-22`) **Coinbase** SEC Lawsuit (`2023-06-06`)	+0.564 +0.706 +0.493 +0.620	-3.58% -0.58% -0.24% -9.28%	**+9.10%** **+5.34%** **+6.36%** **+90.17%**	Multi-year litigation and enterprise lock-in insulate 30-day revenue.
Practitioner Verification	**Apple** M4 Mac Mini Efficiency (`2024-11-12`) **TSMC** US Chip Production (`2024-09-18`) **Amazon** Prime Video Architecture (`2023-05-07`) **Alphabet** Gemini 1.5 Launch (`2024-02-15`)	-0.707 -0.584 +0.043 -0.092	+0.40% +5.30% +0.91% -3.71%	**+14.43%** **+12.63%** **+12.18%** **+4.57%**	Verified price-performance and architectural gains align with positive returns.

4.4 Post-Study Out-of-Sample Verification and Power Bounds

Evaluating N = 28 fresh front-page threads across $28 firms from August 8 to September 25, 2026 confirms out-of-sample consistency with the pooled null. Under our deterministic lexical scorer, the correlation between Net Practitioner Skepticism and 30-day CAR is **r = -0.1279 (p = 0.5166, 95% CI [-0.478, +0.257])** on title plus top four comments and **r = -0.2403 (p = 0.2181, 95% CI [-0.562, +0.146])** on titles alone; under an independent cross-laboratory replication rubric, **r = -0.062 (p = 0.755, 95% Fisher-z CI [-0.43, +0.32])**, with a Day-1 return slope of +0.69 (p = 0.088). At N = 28 (80% power, α = 0.05), the minimum detectable effect size is **MDES r = 0.508** under Fisher-z (0.496 under exact t(26)), ruling out large correlations (|r| > 0.51) while matching the primary null (r = -0.020). Crucially, the post-study cohort also contains **Oracle's September 14, 2026 layoff thread** (*"Oracle's 6am layoff emails hit staff amid new wave of cuts"*, 214 points, 185 comments), which suffered a **-15.42% 30-day CAR** over [t - 1, t + 29] (-13.45% over [t = 0, t + 30]$)—confirming that

layoff reactions depend on distress versus efficiency signaling (Elayan et al., 1998; Hahn & Reyes, 2004) rather than following a mechanical rally rule.

5. Discussion and Conclusion

Across 202 systematic front-page Hacker News threads spanning 54 public technology firms (2022–2026) and 28 out-of-sample threads, grassroots practitioner skepticism does not forecast 30-day abnormal stock returns. Despite high multi-model measurement reliability (ICC(2,4) = 0.9203), both Net Practitioner Skepticism (β = -1.4053%, p = 0.73181) and Hype-Skepticism Divergence (β = -1.5087%, p = 0.69771) lie squarely at zero, and apparent Day-1 return continuation (+0.4531, p = 0.00964) disappears (-0.0053, p = 0.97425) once shared trading session t + 1 is removed.

These findings advance three literatures. First, we extend the online "wisdom of crowds versus noise" debate (Tumarkin & Whitelaw, 2001; Antweiler & Frank, 2004; Tetlock, 2007; Wang et al., 2015; Bartov et al., 2016; Chau et al., 2020; Lachana & Schröder, 2025) from financial message boards to an elite engineering forum. Although software practitioners reduce technical verification costs and create substantial open-source value (Goldfarb & Tucker, 2019; Singh et al., 2026), public engineering critiques of major technology firms are rapidly absorbed by semi-strong efficient equity markets across general corporate announcements (n = 159, β̂ = -1.5639%, p = 0.7243).

Second, we explain why domain-expert sentiment diverges from shareholder valuation. Engineers evaluate corporate actions through product craftsmanship, fair ecosystem pricing, and workplace stability (Dube & Zhu, 2021; Graham et al., 2022), whereas equity investors price residual cash flows under shareholder-value governance (Fligstein & Goldstein, 2022), often welcoming restructurings that unwind technical over-hiring (Hahn & Reyes, 2004; Fedyk & Hodson, 2022) and discounting controversy when enterprise lock-in insulates revenue (Makridis, 2020). Acute product or security failures represent under eight percent of front-page discussions and are offset in pooled panels by restructurings and insulated disputes.

Third, scaling $4.7× from a hand-curated N = 43 pilot to a systematic N = 202 firm-universe corpus illustrates how powered sampling, entity disambiguation (El-Haj et al., 2019), multiple-testing discipline, and out-of-sample verification overcome small-pilot winner's-curse bias (Ioannidis, 2005; Bakker et al., 2012; Szűcs & Ioannidis, 2021; Yang et al., 2023). Because our event-study estimates are observational and focus on 30-day windows using titles and top comments, future work should examine intraday order-book reactions on day t = 0$ and multi-year developer ecosystem attrition. Full datasets, replication code, archived pilot editions, and forensic hashes are provided in the *Technical and Econometric Supplement*.

References

Antweiler, W., & Frank, M. Z. (2004). Is all that talk just noise? The information content of internet stock message boards. *The Journal of Finance*. https://doi.org/10.1111/j.1540-6261.2004.00662.x

Bakker, M., van Dijk, A., & Wicherts, J. M. (2012). The rules of the game called psychological science. *Perspectives on Psychological Science*. https://doi.org/10.1177/1745691612459060

Bartov, E., Faurel, L., & Mohanram, P. S. (2016). Can Twitter help predict firm-level earnings and stock returns? *The Accounting Review*. https://doi.org/10.2308/accr-51865

Chau, M., Lin, C.-Y., & Lin, T.-C. (2020). Wisdom of crowds before the 2007–2009 global financial crisis. *Journal of Financial Stability*. https://doi.org/10.1016/j.jfs.2020.100741

Da, Z., Engelberg, J., & Gao, P. (2011). In search of attention. *The Journal of Finance*. https://doi.org/10.1111/j.1540-6261.2011.01679.x

Das, S. R., Martínez-Jerez, A., & Tufano, P. (2005). eInformation: A clinical study of investor discussion and sentiment. *Financial Management*. https://doi.org/10.1111/j.1755-053x.2005.tb00112.x

Dube, S., & Zhu, C. (2021). The disciplinary effect of social media: Evidence from firms' responses to Glassdoor reviews. *Journal of Accounting Research*. https://doi.org/10.1111/1475-679x.12393

El-Haj, M., Rayson, P., Walker, M., Young, S., & Simaki, V. (2019). In search of meaning: Lessons, resources and next steps for computational analysis of financial discourse. *Journal of Business Finance & Accounting*. https://doi.org/10.1111/jbfa.12378

Elayan, F. A., Swales, G. S., Maris, B. A., & Scott, J. (1998). Market reactions, characteristics, and the effectiveness of corporate layoffs. *Journal of Business Finance & Accounting*. https://doi.org/10.1111/1468-5957.00190

Fedyk, A., & Hodson, J. (2022). Trading on talent: Human capital and firm performance. *European Finance Review*. https://doi.org/10.1093/rof/rfac068

Fligstein, N., & Goldstein, A. (2022). The legacy of shareholder value capitalism. *Annual Review of Sociology*. https://doi.org/10.1146/annurev-soc-030420-120827

Goldfarb, A., & Tucker, C. E. (2019). Digital economics. *Journal of Economic Literature*. https://doi.org/10.1257/jel.20171452

Graham, J. R., Grennan, J., Harvey, C. R., & Rajgopal, S. (2022). Corporate culture: Evidence from the field. *Journal of Financial Economics*. https://doi.org/10.1016/j.jfineco.2022.07.008

Hahn, T., & Reyes, M. G. C. (2004). On the estimation of stock-market reaction to corporate layoff announcements. *Review of Financial Economics*. https://doi.org/10.1016/j.rfe.2003.12.004

Hill, S., & Ready-Campbell, N. (2011). Expert stock picker: The wisdom of (experts in) crowds. *International Journal of Electronic Commerce*. https://doi.org/10.2753/jec1086-4415150304

Huang, A., Zang, A. Y., & Zheng, R. (2014). Evidence on the information content of text in analyst reports. *The Accounting Review*. https://doi.org/10.2308/accr-50833

Ioannidis, J. P. A. (2005). Why most published research findings are false. *CHANCE*. https://doi.org/10.1080/09332480.2005.10722754

Lachana, I., & Schröder, D. (2025). Investor sentiment and stock returns: Wisdom of crowds or power of words? Evidence from Seeking Alpha and Wall Street Journal. *Journal of Financial Markets*. https://doi.org/10.1016/j.finmar.2025.100970

Makridis, C. A. (2020). Do data breaches damage reputation? Evidence from 43 companies between 2002 and 2018. *Journal of Cybersecurity*. https://doi.org/10.1093/cybsec/tyab021

Singh, S. P., Ramasubbu, N., Kemerer, C. F., & Hydari, M. Z. (2026). Profitability of open-source software product development. *Information Systems Research*. https://doi.org/10.1287/isre.2025.1821

Szűcs, D., & Ioannidis, J. P. A. (2021). Empirical assessment of published effect sizes and power in the recent cognitive neuroscience and psychology literature. *PLoS Biology*. https://doi.org/10.1371/journal.pbio.3001151

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*. https://doi.org/10.1111/j.1540-6261.2007.01232.x

Tumarkin, R., & Whitelaw, R. (2001). News or noise? Internet postings and stock prices. *Financial Analysts Journal*. https://doi.org/10.2469/faj.v57.n3.2449

Wang, G., Wang, T., Wang, B., Sambasivan, D., Zhang, Z., Zheng, H., & Zhao, B. Y. (2015). Crowds on Wall Street. *ACM Conference on Computer Supported Cooperative Work (CSCW)*. https://doi.org/10.1145/2675133.2675144

Yang, Y., Sánchez-Tójar, A., O'Dea, R. E., Noble, D. W. A., Koricheva, J., Jennions, M. D., Parker, T. H., Lagisz, M., & Nakagawa, S. (2023). Publication bias impacts on effect size, statistical power, and magnitude (Type M) and sign (Type S) errors in ecology and evolutionary biology. *BMC Biology*. https://doi.org/10.1186/s12915-022-01485-y