Commoditize Your Complement: Strategic Open-Source Releases and the Cannibalization of Rival Firms

Peter Young Won *

June 2026

ABSTRACT

Why do firms release open-source software (OSS) for free? I show that platforms strategically release OSS to commoditize complementary product markets, eroding rivals' proprietary rents while expanding demand for their own products. Using the staggered release of GitHub repositories, I find that complement rivals lose 8.2% of sales, cut R&D by 13.1%, and reallocate engineering talent to the platform, whereas direct substitutes show no measurable reaction. The platforms benefit in turn: their core-segment sales rise about 42% more than non-core segments, lifting sales and profitability. The stock market reacts the same way: announcement-window cumulative abnormal returns (CAR) capitalize the transfer in real time, with the releasing platforms adding roughly $42 billion in market value while complement rivals lose about $4 billion. Free OSS emerges as a non-priced competitive weapon beyond the reach of traditional antitrust.

JEL Classification: L12, L17, O31, O32, G14, G32

Keywords: Open-Source Software, Commoditization, GitHub Ecosystem, Innovation

Acknowledgements and Disclosure

*Won: peter.young.won@eccles.utah.edu.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The author declares no financial relationships or other potential conflicts of interest relevant to this research. The analysis uses publicly available data (the GitHub REST API, GH Archive, the USPTO via PatentsView, and the Hoberg-Phillips Data Library) together with Compustat and CRSP accessed under an institutional Wharton Research Data Services (WRDS) subscription. As the sole author, I had full access to all data used in the study and take responsibility for the integrity of the data and the accuracy of the analysis.

1 Introduction

Why do firms release open-source software for free? Over the past decade, Alphabet, Meta, Microsoft, and a handful of other dominant technology firms have published billions of lines of source code to the public at no charge. Google gave away TensorFlow, the framework that runs its own search and advertising models, and Meta did the same with PyTorch and React. Each release hands rivals software at no cost that took years and hundreds of engineers to build. A growing literature treats these releases as a public good that raises productivity for the firms that adopt them (Nagle, 2019; Hoffmann et al., 2024). Platform theory points to a second possibility: by giving away a high-quality product that complements its own core business, a dominant firm makes the complement cheap, pulls demand toward the core where it earns its profits, and undercuts the rivals that sold the now-free product (Parker and Van Alstyne, 2005; Economides and Katsamakas, 2006; Gawer and Henderson, 2007). Whether this commoditize-your-complement mechanism redistributes rents away from incumbents rather than lifting all firms together has not been measured. I provide the first large-sample evidence.

A historical precedent makes the mechanism concrete. Gurley (2026) observes that the outcome of cloud computing's first two decades is not that Amazon Web Services came to dominate, but that IBM, Hewlett-Packard, and Intel, the owners of the underlying stack, did not: open-source commoditization of that stack dissolved their supplier power, and the operator captured the rents instead. This paper documents the same dynamic arising by deliberate corporate strategy, with dominant incumbents commoditizing complementary layers through their own releases.

Measuring the competitive effect on incumbent rivals requires three ingredients at once: a firm-level measure of exposure to each release, a set of release shocks comparable in economic magnitude, and a credible counterfactual for the same firms absent the release. The existing literature on open source supplies none of the three, having concentrated on the releasing firm or on aggregate productivity (Nagle, 2019; Hoffmann et al., 2024). I assemble all three. I combine repository-level developer activity from GitHub Archive with firm-level Compustat outcomes and USPTO patent records, I retain roughly 1,700 mechanically eligible releases and weight each by a market-value intensity measure following Kogan et al. (2017), and I map each release to its product-market peers using text-based network industry classifications (Hoberg et al., 2014; Hoberg and Phillips, 2016).

A stacked difference-in-differences estimator, weighted by this market-implied release intensity, then traces how each cohort reshapes peers' revenues, costs, innovation, and labor.

The rival response unfolds in three steps. First, I find that revenue contracts: treated peers' sales fall by 8.2 percent relative to matched controls in the post-release window. Facing this compression, the same firms cut direct operating costs by substituting the now-free open-source infrastructure for internal proprietary inputs, which produces a short-run lift in profitability that masks the underlying erosion.

Second, I find that this cost adjustment operates through the engineering labor market. Using developer engagement logs from GitHub across the full set of releases, I observe that rival engineering capital flows onto the focal platform's projects: active development staffing at exposed peers rises by about 15 percent and total contributions per developer by about 37 percent. As rivals adopt the free release, their engineers spend a growing share of their effort building on the focal firm's commons rather than on proprietary work (Gupta et al., 2024).

Third, I show that this reallocation of engineering effort carries a long-run innovation cost. Research and development scaled by assets falls by 13.1 percent over the same window, the average age of treated firms' patent portfolios rises, and the share of frontier exploratory patents declines.

These averages conceal a sharp asymmetry. I partition treated peers by Cooperative Patent Classification overlap and pre-event return co-movement into substitute peers, which compete head to head with the focal firm, and complement peers, which operate one layer adjacent to it. I find that the revenue, innovation, and labor effects load almost entirely on complement peers, with substitutes largely insulated on sales; substitutes do cut direct costs, though by less than complements, and their market reaction is small and second-order. The equity market prices the same split in real time. Over the $[0, +3]$ window around the release, value-weighted cumulative abnormal returns are $+0.32$ percent for releasing platforms, -0.26 percent for complement peers, and a small $+0.09$ percent for substitutes. Capitalized per unique firm, the market marks up the releasing platforms by roughly \$42 billion in aggregate while marking complement rivals down, leaving substitutes close to unchanged. The product-market and asset-price evidence line up on the same firms, which rejects diffuse spillover and identifies the release as a targeted shock to the platform's complementary layer.

The release does not only subtract from rivals; it adds to the firm that issues it. The same re-

leases that compress complement rivals expand the releasing firm's own core business. Within firm and year, the platform's core-segment sales rise about 42 percent more than its non-core-segment sales after the release, and firm-level sales and profitability rise relative to matched controls. The strategy is therefore two-sided: a single act that cannibalizes complementary rivals while channeling the freed demand into the platform's own profit center.

The identification rests on three design choices. Peer assignment is anchored to the lagged TNIC product-market network, so a release cannot rewrite the peer set it is supposed to affect. Cohorts are weighted by Kogan release intensity, so the estimate reflects the economically influential events rather than the long tail of trivial ones. Controls are one-to-one propensity-score matches drawn from outside the focal firm's network (Baker et al., 2022). Under this design, the post-release dynamics are not pre-existing industry trends but the direct consequence of the release.

This paper makes three contributions. First, it provides the first large-sample, rival-level test of the commoditize-your-complement logic (Parker and Van Alstyne, 2005; Economides and Katsamakas, 2006; Gawer and Henderson, 2007). The empirical literature on open source has measured aggregate productivity gains (Nagle, 2019), the financing of complementary startups (Conti et al., 2025), and the private value captured by the contributing firm (Hoffmann et al., 2024; Zhao, 2025), but the rival whose proprietary market is commoditized has not been the unit of analysis. I locate the strategy at the rival level, I translate the market reaction into dollar-valued wealth transfers, I use a continuous text-based network measure to identify the exposed peer set and then a technology-overlap and return-co-movement partition to separate the harmed peers from the insulated ones, and I trace the freed demand to the releasing firm's own core segments.

Second, I extend the resource-based view and the open-innovation literature to the involuntary case. That literature argues that voluntary surrender of proprietary control can be rational for the disclosing firm (Alexy et al., 2018, 2013; Chesbrough and Crowther, 2006). I show the other side of the transaction: when a dominant platform releases a complement at zero price, exposed rivals not only watch the value of their proprietary stack erode but also reallocate engineering effort toward the platform's commons. Labor reallocation and innovation decay therefore emerge as the hidden costs of defensive open-source adoption (Mkrtchyan et al., 2025; Heath et al., 2023; Gupta et al., 2024), widening the set of margins along which competitive harm travels in digital product markets.

Third, the findings speak to a regulatory gap. Traditional antitrust screens platform conduct for predatory pricing, output restriction, or explicit exclusion in the core market. A free release of a complementary technology satisfies none of these doctrines, yet I show that it extracts rents from rivals, slows their independent innovation, and absorbs their engineering capacity. As data and digital infrastructure increasingly function as networked competitive assets (Bian et al., 2025), an antitrust framework that price-screens only the focal product will miss the non-priced competitive instrument documented here.

The remainder of the paper is organized as follows. Section 2 positions the paper relative to five strands of prior work. Section 3 develops a simple multi-segment model that delivers the asymmetric sign prediction tested in the main analysis. Section 4 describes the data and the stacked difference-in-differences identification strategy. Section 5 reports the baseline rival-side findings on revenue, cost, labor, innovation, and product-market dynamics. Section 6 introduces the substitute–complement partition, shows that the average effects load on complement peers, presents the contemporaneous wealth-transfer evidence from cumulative abnormal returns, and documents the releasing firm's own core-demand expansion. Section 7 concludes.

2 Related Literature

This paper draws on five lines of prior research: the economics of free complements, two-sided markets, and network effects; the theory of platform competition and complementor response; empirical work on strategic open-source release and selective revealing; the measurement of innovation spillover and product-market rivalry; and the resource-based view (RBV) of the firm. Each supplies part of the vocabulary for a commoditize-your-complement strategy, but none predicts the sign asymmetry this paper shows, in which substitute peers are insulated while complement peers are harmed.

The idea that a firm might accept lower rents in one market to raise them in another, articulated as industry folklore by Spolsky (2002), is formalized in the economics of complements, two-sided markets, and network effects. The pricing of complements dates to Cournot (1838), who showed that separate owners of two complementary goods set prices above the level a single owner of both would choose, so whoever controls one layer gains when the price of the complement falls; Varian

(2004) develops this systems-of-complements logic for information-technology markets, where a firm routinely lowers the price of a complement to expand demand for its profit center. Parker and Van Alstyne (2005) show that an information good with strong cross-network externalities on a paid core is optimally priced at zero, Rochet and Tirole (2003) and Armstrong (2006) characterize when a platform optimally subsidizes one side to expand the other, and Katz and Shapiro (1985), Farrell and Saloner (1985), and Shapiro and Varian (1999) show how network benefits, standardization, and zero marginal cost combined with lock-in make a zero or even negative price equilibrium-consistent for digital goods. Economides and Katsamakas (2006) extends the logic to the choice between a proprietary and an open-source regime, and the closest empirical antecedent, Casadesus-Masanell and Ghemawat (2006), models an open-source good coexisting with a proprietary incumbent under strong demand-side learning. These models explain why a zero price is optimal, but they treat the open-source good as arising bottom-up from a community rather than as a top-down release by a multi-segment incumbent that internalizes both the rent loss in the commoditized segment and the rent gain in its core, and they say little about which rivals bear the loss. That internalization is what makes rival displacement an equilibrium outcome rather than a byproduct of community competition.

A second strand turns from the focal firm's incentive to the rivals and complementors that a platform's choices affect. Gawer and Henderson (2007) describes the managerial practice of cultivating complementary ecosystems, while Eisenmann et al. (2011) on platform envelopment and Wen and Zhu (2019), Cennamo et al. (2018), and Wen et al. (2016) on the threat of platform-owner entry show how an aggressive platform reshapes complementor behavior and investment incentives. These frameworks treat the affected rivals as undifferentiated, however, and they typically have the focal firm enter the adjacent market, whereas an open-source release commoditizes that market without entry. The closest structural antecedent of this paper's peer classification is Adner and Kapoor (2010), who distinguish upstream components from downstream complements and show that performance across technology generations depends on which side of that divide a firm occupies; this paper operationalizes that distinction as the Substitute–Complement peer partition and shows it is, empirically, the primary axis along which release effects sort.

A more empirically grounded literature documents the strategic motives behind open-source release at the firm and project level: Lerner and Tirole (2002) on why firms contribute at all,

Henkel (2006) on selective revealing in embedded Linux, Fosfuri et al. (2008) on commercializing releases as a product-market strategy, Alexy et al. (2013) and Alexy et al. (2018) on when surrendering proprietary control benefits the revealing firm, and Casadesus-Masanell and Llanes (2011) on profitable mixed open-and-proprietary strategies. A recent wave uses firm-level data to estimate the private value of adopting or contributing to open source (Nagle, 2019; Hoffmann et al., 2024; Zhao, 2025; Nagaraj and Piezunka, 2024; Haese and Peukert, 2025) and how participation interacts with corporate investment and financing (Conti et al., 2025; Emery et al., 2025; Coleman et al., 2025). This work has concentrated on the releasing or adopting firm and has relied largely on cross-sectional adoption variation or case studies; the rival whose proprietary business is commoditized has rarely been the unit of analysis. This paper instead treats approximately 1,700 economically consequential releases as staggered shocks and places the rival firm's financial, innovation, and labor outcomes at the center of measurement.

A large literature separates knowledge spillover from product-market rivalry (Jaffe, 1986; Aghion et al., 2005; Makaew and Maksimovic, 2020; Baker et al., 2023), with Bloom et al. (2013) finding that spillovers dominate in aggregate. This paper is its mirror image: the release is itself the spillover, so the net rival effect splits by product-market position rather than averaging to a single sign, a pattern that prior single-shock studies (Seamans and Zhu, 2014; Boudreau, 2010) do not isolate. The resource-based view (Barney, 1991) and its extension to voluntary openness (Alexy et al., 2018; Chesbrough and Crowther, 2006) say little about the involuntary case studied here, in which a third party's release erodes a rival's resource rarity and forces adjustments along its innovation and labor margins (Benner and Tushman, 2002; Manso, 2011; Mkrtchyan et al., 2025; Lin and Maruping, 2022; Heath et al., 2023; Gupta et al., 2024).

Taken together, these literatures supply the theoretical vocabulary but not the empirical architecture to identify a commoditize-your-complement strategy as it operates on rivals: they explain the focal firm's incentive, the optimal zero price, the releasing firm's motives, and aggregate spillover effects, but none identifies which rivals lose. This paper combines a large-sample event design, a rival-level outcome panel, and a Substitute–Complement partition grounded in the upstream–downstream distinction of Adner and Kapoor (2010), and traces the tradeoff in which a focal firm gains in its profit-center segment by commoditizing a complementary one, transferring wealth from rivals whose revenues reside in that segment while leaving rivals in adjacent segments

untouched or even better off.

3 Theoretical Framework and Hypothesis Development

This section sets out a deliberately simple model that organizes the empirical analysis and states the five hypotheses it generates. The mechanism, in the vocabulary of Porter's Five Forces, is the elimination of upstream supplier power through commoditization of the complement layer, with the rent then accruing to the platform operator rather than the IP owner. Gurley (2026) describes Amazon Web Services as the canonical historical instance: exogenous open-source commoditization of the operating-system and middleware stack left the operator position empty for Amazon to occupy, while IBM, Hewlett-Packard, and Intel, the firms that should have extracted the supplier rents, were left with no surface on which to charge. The framework below makes the same logic endogenous, with incumbents financing the commoditization themselves out of profit-center rents.

This structure builds on the two-sided cross-subsidy model of Parker and Van Alstyne (2005), in which a firm facing cross-market network effects optimally prices a complementary good at zero to raise demand for its paid core. The contribution here is to specialize that logic to a multi-segment incumbent and to extend it from the focal firm's own pricing problem to the asymmetric response of exposed rivals. The simple model, comprising the multi-segment profit function, the four-channel decomposition of the release decision, and the derivations of Propositions 1 and 2, is developed in full in Appendix A. Figure 1 summarizes the framework's causal logic, from the release decision through the asymmetric peer split to the resulting wealth transfer.

The analytical primitive is the multi-segment firm, operating across three logically distinct product markets. Segment A is the profit center that generates the firm's protected rents, such as advertising inventory for Meta, search advertising for Alphabet, and enterprise cloud for Microsoft. Segment B is a separate market the focal firm also occupies but that a release in C leaves unaffected; it plays no role in the peer cross-section, which turns on segments A and C. Segment C is a complement market in which higher consumption raises the marginal value of A through a cross-market demand linkage. In segment C the firm's optimal posture is commoditization rather than rent extraction, because cheaper C-goods expand demand for the profitable A-segment. A strategic open-source release is exactly this act: the focal firm gives away a high-quality asset in segment C

at a consumer price of zero, financed out of segment A's rents. The open-source license makes the released version difficult to withdraw, while the cross-subsidy itself comes from the firm's multi-segment economics rather than from the license; together these separate the act from an ordinary price cut and render the standard predatory-pricing hold-up logic inapplicable.

Releasing is optimal when the cross-market subsidy to segment A, together with the option value of setting the technical standard and of co-opting outside developer effort, dominates the direct loss of segment-C rent. This condition holds for incumbents whose profit base is dominated by a large, protected A-segment, the firms with a subsidy pool deep enough to finance a permanent giveaway, and fails for firms whose profits are themselves concentrated in segment C. Because a release with negative net present value would not be announced, the focal firm's announcement-window equity return is non-negative in expectation.

Proposition 1 (Release under Profit-Center Dominance): *A focal firm releases a strategic open-source asset in its complement market when its profit-center rents are large relative to its complement-segment rents and the cross-market demand linkage is sufficiently strong. The release is then optimal even though it forgoes the firm's own complement-segment revenue.*

Hypothesis 1 (Focal-Firm Valuation Gain): *A strategic open-source release generates a non-negative cumulative abnormal return, on average, on the focal firm's equity in a narrow window around the release.*

The framework's sharpest prediction concerns the cross-section of peer firms, and it is pinned down by where each peer's revenue lives across the three segments. A *substitute peer* earns its revenue in the focal firm's profit center A; a *complement peer* earns it in segment C, the very market the release commoditizes. For a complement peer, the release directly compresses pricing power: consumers facing a free, high-quality alternative abandon its proprietary C-product, and its revenue base contracts with no offsetting revenue in A-adjacent markets. For a substitute peer, there is no direct C-segment compression. The expanded complement market does raise demand for the substitute peer's own profit-center revenue, since segment C complements segment A for every firm selling there, not only for the focal firm. The focal firm, however, is dominant in segment A and is the direct beneficiary of the complement it released, so it absorbs the bulk of this demand expansion, leaving the substitute peer a second-order residual bounded above by the focal firm's

own absorption, weakly positive or statistically indistinguishable from zero.

Proposition 2 (Asymmetric Response): *Announcement of a strategic open-source release generates a strongly negative cumulative abnormal return on complement peers and a weakly positive or null return on substitute peers. The complement-peer loss is larger the more the peer's revenue is concentrated in the commoditized segment, while the substitute-peer effect is bounded above by the focal firm's own absorption of the expanded complement-market demand.*

Hypothesis 2 (Asymmetric Peer Valuation): *A strategic open-source release generates a strongly negative cumulative abnormal return on complement peers and a weakly positive or null response on substitute peers that is second-order relative to the complement-peer loss; the asymmetry follows from peer portfolio composition and assumes nothing about the focal firm's targeting intent.*

Hypothesis 2 is the framework's central prediction, and the one that distinguishes it from competing explanations. No alternative account of the release reproduces this pattern. A quality-signaling story implies a uniform sign across peers, since all observers update the same prior about the focal firm. Corporate philanthropy implies a valuation shock at the focal firm with no systematic peer effect, and network bootstrapping benefits complementors through wider adoption rather than harming them. Bertrand predation is the closest competing account. The focal firm's positive announcement return is hard to square with classic predation, in which the predator sacrifices current profit to discipline rivals, but a positive return alone does not rule out an exclusionary strategy whose expected future gains investors capitalize; the decisive distinction is the cross-peer pattern, since predation does not generate the substitute-versus-complement asymmetry that Hypothesis 2 predicts. Appendix A develops each case. Section 6.1 details how the latent substitute and complement types are projected onto observable pre-event signals.

The negative valuation response on complement peers propagates into firm behavior across three real-economy margins. Facing a credible, permanent commoditization of its C-segment product, the complement peer loses pricing power and its revenue contracts; at the same time it adopts the free release as an input substitute, cutting its cost of goods sold fast enough to raise asset-scaled profitability in the short run. The peer's engineering labor is redirected toward maintaining compatibility with the focal firm's evolving standard, measurable as contribution activity on the focal

firm's public repositories. As that effort migrates to the focal firm's commons and the marginal return to proprietary *C*-innovation falls, internal exploratory research becomes less attractive than exploitation and its patent portfolio ages.

Hypothesis 3 (Financial Compression and Short-Run Profitability Expansion): *Complement peers exhibit declining sales and declining cost of goods sold, with the cost contraction outpacing the revenue decline so that asset-scaled profitability rises in the short run.*

Hypothesis 4 (Labor Co-optation): *Complement peer developers increase active contribution (commits, pull requests, and issues) to the focal firm's open-source repositories.*

Hypothesis 5 (Innovation Retreat): *Complement peers reduce internal R&D expenditure and experience aging of their exploratory patent portfolios.*

Hypotheses 3 through 5 together characterize the complement peer's adjustment path. On the producer side the framework predicts a focal-firm gain, a complement-peer loss, and an approximate substitute-peer null; the net effect on consumer surplus is left open, and the paper documents this asymmetric repricing as a positive economic fact. Appendix A develops the simple model and all derivations.

4 Data and Sample Construction

This study constructs a novel, multi-source panel designed to isolate the strategic impact of open-source software releases by dominant technology platforms on the financial performance, innovation trajectories, and human-capital allocation of product-market rivals. The data construction integrates four independently sourced components. First, I collect repository-level metadata and release histories from the GitHub REST API.[1] Second, I extract daily developer-activity logs from the GH Archive on Google BigQuery.[2] Third, I incorporate firm-year accounting, market, and patent data from Compustat, CRSP, and the USPTO. Finally, I map the competitive landscape us-

[1]The official documentation for the GitHub REST API is available at `https://docs.github.com/en/rest`.

[2]Historical GitHub event logs are accessed via the GH Archive public dataset hosted on Google BigQuery, detailed at `https://www.gharchive.org/`.

ing text-based product-market network measures from the Hoberg-Phillips TNIC database.[3] The primary empirical contribution of this paper lies in the integration of these components. Rather than selecting a discretionary set of high-profile releases, I retain every mechanically eligible release and use a market-value-based intensity measure adapted from Kogan et al. (2017) as the analytic weight in the stacked difference-in-differences estimator throughout the analysis, so that the economically consequential releases drive the estimate without the event set itself being chosen on that intensity.

4.1 Identifying Strategic Open-Source Releases: The Kogan Intensity Measure

The construction of the event sample makes two distinct methodological choices, one about which releases enter the panel and one about how each included release is allowed to influence the estimate. Treating these as separate decisions, rather than collapsing them into a single ex-post selection rule, is what addresses the two competing concerns about the event sample: selection on outcome-correlated characteristics, and signal attenuation from economically marginal releases.

On the sample side, I depart from the prevailing practice in the corporate open-source literature of restricting the event set to a discretionary list of high-profile releases identified through cumulative community stars, fork counts, or downstream adoption. Each of those metrics is realized over years of community uptake and is therefore mechanically correlated with the same competitive outcomes the analysis seeks to estimate; using them as a selection filter embeds look-ahead bias with respect to the response the research design is intended to measure. To eliminate this circularity, I include every open-source repository released by one of the focal technology firms between January 2015 and December 2024 for which CRSP coverage of the releasing firm and pre-event market capitalization are available. The single inclusion filter is therefore mechanical and ex-ante, not discretionary, and the resulting universe of approximately 1,700 release cohorts is fixed before any outcome is measured.[4] Each cohort's event date is the repository's earliest

[3]The product market data are publicly available at the Hoberg-Phillips Data Library: `https://hobergphillips.tuck.dartmouth.edu/`.

[4]The headline results are robust to further restricting the sample to the approximately 1,300 economically substantial releases, after excluding clearly minor ones such as bug fixes, documentation or dependency updates, tutorials, and sample code. Minor and substantial releases are separated by a multi-model large-language-model screen applied

open-source release recorded within the 2015–2024 sample window, and the abnormal-return and difference-in-differences windows below are anchored to that date.[5]

On the weighting side, I assign to each release a continuous, market-implied intensity weight constructed entirely from pre-event and contemporaneous stock-market information. For repository r released by firm $f(r)$, the cohort-level Kogan intensity is

$$\kappa_r \;=\; \mathrm{CAR}^{\mathrm{mkt}}_{r,[-1,+1]} \;\times\; \mathrm{MktCap}_{f(r),-2},$$

where $\mathrm{CAR}^{\mathrm{mkt}}_{r,[-1,+1]}$ is the cumulative market-adjusted abnormal return of $f(r)$ over a three-trading-day window centered on the release announcement, computed as the firm's daily return in excess of the CRSP value-weighted market return, and $\mathrm{MktCap}_{f(r),-2}$ is the firm's market capitalization two trading days before the release, the close immediately preceding the CAR window.[6] The construction directly parallels the patent-value measure in Kogan et al. (2017), in which the dollar-value impact of a patent grant is recovered from the issuing firm's narrow-window abnormal return multiplied by its pre-event market capitalization. Applying the same logic to open-source releases converts an unranked technological event into a continuous, market-implied dollar-magnitude proxy. I use the absolute value $|\kappa_r|$ as the analytic weight throughout the stacked difference-in-differences specification, since the focus is on the structural magnitude of the competitive shock rather than on the releasing firm's directional wealth gain or loss.

The two choices address two distinct concerns that the prior literature typically conflates. Fixing the included event set at the mechanical 1,700-cohort universe rules out selection on outcome-correlated characteristics: nothing about a release's eventual community success or downstream adoption gates its presence in the panel. Weighting by $|\kappa_r|$, in turn, rules out attenuation bias

to each repository's description, topics, and language metadata, not to any outcome- or popularity-based measure, so the screen preserves the ex-ante character of the event set.

[5]For long-running repositories whose original public launch predates the window, this is the first qualifying release observed inside the window rather than the original launch; for example, the golang/go cohort is dated to the Go 1.4.1 release in January 2015, not Go's 2009 debut. Such cases concentrate in early 2015 and are a small share of the cohort universe. Collapsing the few owner-name capitalization variants that appear in the raw extract is deferred to a subsequent revision and does not affect the headline estimates.

[6]The headline estimates are robust to how the intensity weight is built. Reconstructing $|\kappa_r|$ from a Fama-French three-factor abnormal return augmented with the Carhart momentum factor, or from the original Kogan et al. (2017) $[0,+2]$ event window in place of the centered $[-1,+1]$ window, preserves the sign and significance of all four real-outcome estimates. The Kogan-weighted sales coefficient, for instance, is -0.052 under the four-factor weight and -0.056 under the $[0,+2]$ window, against -0.059 in the baseline, all significant at the 1% level.

from economically marginal releases: each cohort is weighted by its market-implied dollar magnitude rather than receiving equal influence in the estimate, so the headline coefficient reflects the response to the economically consequential events rather than the average of consequential and inconsequential ones. Internet Appendix B.1 confirms that these two concerns are empirically separated as the design intends. A saturation curve that re-estimates the headline coefficient on Top-K subsets of cohorts ranked by $|\kappa_r|$ shows that the $|\kappa_r|$-weighted estimate is essentially invariant to the sample cut, while a complementary trimming-from-bottom exercise shows that the unweighted estimate exhibits the textbook attenuation pattern the weighting is designed to correct.

Figure 2 plots the annual distribution of the 1,700 release cohorts. Consequential releases concentrate between 2015 and 2017, the period when leading technology firms decisively pivoted toward open-source leadership, followed by a meaningful secondary wave in 2019 and 2020. This continued flow provides the staggered temporal variation needed for the stacked difference-in-differences estimator.

[Insert Figure 2 Here]

Table 1 reports descriptive evidence that the intensity measure captures meaningful cross-cohort variation in economic magnitude. Panel A documents a wide dispersion across κ_r quartiles, ranging from an average loss above 16 billion dollars in the bottom quartile to a gain above 19 billion dollars in the top quartile. Average developer engagement measured by cumulative GitHub stars remains uniformly high across quartiles, indicating that all cohorts in the universe represent substantive technological events; the intensity weight differentiates them on market-implied dollar magnitude rather than on adoption per se. Panel B shows the extreme concentration of consequential deployment among a small number of dominant platforms: Microsoft, Alphabet, and Meta together account for the majority of repositories and of cumulative intensity, motivating the focus on those firms as the primary drivers of platform-level commoditization.

[Insert Table 1 Here]

Figure 3 plots the cohort-level distribution of $|\kappa_r|$. The raw distribution is heavily right-skewed, in line with standard distributional properties of innovation and patent-value measures; the log-transformed distribution closely approximates a normal curve, neither bimodal nor disproportionately driven by a small cluster of outliers. This continuous log-normal shape supports the use of

$|\kappa_r|$ as a continuous analytic weight in the stacked difference-in-differences specification, ensuring that the empirical estimates are weighted by the underlying economic magnitude of the release rather than treating every repository as an equal shock.

[Insert Figure 3 Here]

Two additional features of the intensity measure are worth stating explicitly. First, κ_r is computed entirely from pre-event and contemporaneous market data, ensuring that it remains uncontaminated by the subsequent product-market adjustments, patenting responses, or labor reallocations evaluated as outcomes in this study. Second, relative to the original Kogan et al. (2017) construction, this adaptation does not subject the release announcement window to an exhaustive news-filtering procedure, so a marginal fraction of the intensity values may absorb concurrent unrelated corporate news. The saturation and trimming exercises in Internet Appendix B.1 indicate that this residual noise exerts negligible influence on the headline estimate, consistent with the inverse-variance logic that motivates the weighting choice in the first place.

4.2 GitHub Repository and Developer Data

To construct the universe of open-source releases and observe developer-side labor reallocation at a daily frequency, I assemble a granular dataset utilizing the GitHub REST API and Google BigQuery. This collection and refinement pipeline is explicitly designed to isolate externally sourced engagement from internal corporate development.

I first identify the focal technology firms whose corporate accounts drive the modern open-source ecosystem, including large platform firms such as Microsoft, Alphabet, Meta, and Apple, alongside infrastructure and semiconductor firms such as Nvidia, Adobe, and Oracle. Focusing on this concentrated set of firms is economically motivated. The open-source landscape is heavily skewed toward the top, where foundational platforms, standards-setting libraries, and high-impact infrastructure projects are overwhelmingly authored by a relatively small set of large technology firms. I map each focal firm to its authenticated organizational accounts, such as google, tensorflow, and microsoft. To prevent the inclusion of unaffiliated mirrors or community forks, I verify that each organizational account is operated by the focal firm itself, confirming that the account's owners and core organization members post from the firm's verified corporate domains, such as

microsoft.com and oracle.com. Following this account-level verification, I extract metadata for all public repositories owned by these accounts released between January 2015 and December 2024. This account-ownership check, which fixes the set of release events, is distinct from the developer-level employment mapping used to measure labor reallocation, which links individual contributors to their employing firms and is described below. Applying the mechanical inclusion filter described in Section 4.1 yields the approximately 1,700 eligible cohorts utilized throughout the main analysis, each subsequently entering the panel under its $|\kappa_r|$ analytic weight. Internet Appendix Table IA.1 lists representative examples of these releases across the focal firms.

For the developer-level analysis of labor reallocation in Section 5.3, I resolve contributor identity to the employing firm across the full set of approximately 1,700 release cohorts, processing the complete GitHub Archive event history on BigQuery and linking each contributor to a Compustat GVKEY through the firm's verified corporate domains. The headline labor estimates in Table 5 therefore use the same full-cohort universe as the rest of the analysis. As a sharper-measurement robustness check, I additionally isolate the top 100 cohorts ranked by Kogan release intensity, for which the per-developer panel is densest; this subset selects on the release's announcement-window magnitude, which is predetermined relative to the labor outcomes rather than to the release itself.

Using the GH Archive on BigQuery, I extract all event-level activity records for the target repositories, including pushes, pull requests, issue submissions, stars, and watches. I classify these records into two economically distinct categories. Active contributions, comprising pushes, pull requests, and issues, capture direct developer effort and committed human capital. Passive contributions, consisting of stars and watches, capture broader community attention and technological adoption. This bifurcation maps directly onto the theoretical distinction between labor reallocation, which serves as the core mechanism of commoditization in this study, and adoption spillovers, which shape longer-run platform network effects.

Linking individual GitHub developers to their employing firms constitutes the central measurement challenge in the labor reallocation analysis. I extract email domains and self-reported company names from developer profiles and commit logs, matching them to the verified corporate domains of publicly traded firms. These matched domains subsequently resolve to standard Compustat GVKEY identifiers. To ensure identification accuracy, I strictly drop generic personal

email providers, such as gmail.com and yahoo.com, from the matching process. This econometric choice deliberately privileges a low Type I error rate at the cost of a higher Type II error rate. While excluding generic domains fails to capture corporate developers contributing via personal accounts, it essentially eliminates the risk of misattributing independent or hobbyist developer activity to a specific public firm. Consequently, the resulting developer-to-firm panel represents a conservative lower bound on actual corporate engagement. Any statistically significant treatment effect documented in this analysis thus survives substantial measurement-error attenuation toward zero.

Finally, to guarantee that the measured engagement reflects genuine external labor reallocation rather than the focal firm's internal research and development, I impose a strict self-loop filter. All GitHub activity generated by developers affiliated with the focal releasing firm is entirely removed before aggregation. For instance, contributions by Microsoft employees to a Microsoft-owned repository are excluded from the peer engagement metrics. The filtered panel therefore cleanly captures the external labor mobilization of competing firms redirecting their human capital toward the newly established open-source infrastructure.

4.3 Firm Financials, Market Valuation, and Innovation Outcomes

To evaluate the financial, operational, and innovation consequences of an open-source release, I construct a firm-year panel using accounting data from Compustat, market data from the Center for Research in Security Prices (CRSP) via the CRSP-Compustat Merged database, and patent data from the United States Patent and Trademark Office (USPTO). The sample spans fiscal years 2010 through 2024 to accommodate the pre- and post-event windows around the eligible open-source releases, subject to the data-window truncation noted in Section 4.6. Following standard practice in empirical corporate finance, I exclude financial institutions and regulated utilities, as their capital structures are not comparable to those of the technology and industrial firms that populate the peer groups in this study.

From this integrated panel, I compute the primary outcome variables utilized throughout the analysis. To evaluate revenue impact and profitability, I construct measures for Sales, Cost of Goods Sold (COGS), and Return on Assets (ROA), each scaled by total assets to adjust for firm

size. To measure innovation outputs, I calculate Research and Development (R&D) expenditure scaled by total assets. Furthermore, following Manso (2011) and Balsmeier et al. (2017), I classify patents as either exploratory or exploitative using Cooperative Patent Classification (CPC) codes. Exploratory patents are defined as those filed in CPC classes new to the patenting firm itself, whereas exploitative patents are those filed in that same firm's incumbent classes. Based on this classification, I compute the average age of the firm's overall patent portfolio and the average age of its exploratory patents. Finally, the firm-level control variables utilized in the baseline regressions, specifically Firm Size, Market-to-Book Ratio, and Leverage, are constructed following the conventions established by Fama and French (1992), Frank and Goyal (2009), and Roberts and Whited (2013).

4.4 Text-Based Product Market Networks

Identifying the firms exposed to a release requires a firm-specific, bilateral measure of which rivals occupy the same product market, which standard industry codes do not supply. Fixed SIC and NAICS classifications are too coarse and too static for this, and they assign multi-segment technology firms to categories that obscure their actual product overlap. I therefore use the Text-based Network Industry Classifications (TNIC) of Hoberg and Phillips (2016), which are built from firms' 10-K product descriptions, defined at the firm-pair level, and updated annually, and which pin down the peer set around each release.[7] Unlike static industry codes, the TNIC methodology computes continuous pairwise similarity scores based on the cosine similarity of firms' 10-K product descriptions. This yields a dynamic map of competition that is particularly well-suited to the fast-evolving technology sector. I employ TNIC similarity scores to identify the product-market peers of each focal releasing firm. Internet Appendix Table IA.2 illustrates the face validity of this assignment, listing the highest-similarity TNIC peers for the major focal firms.

4.5 Treatment Assignment, Matching, and Cohort Construction

For each eligible open-source release r by a focal firm $f(r)$, treated peers are defined as all Compustat firms exhibiting strictly positive TNIC similarity to $f(r)$ in the fiscal year immediately

[7] The data are available at the Hoberg-Phillips Data Library, `https://hobergphillips.tuck.dartmouth.edu/`.

preceding the release, $t-1$.[8] Utilizing the lagged product-space network deliberately eliminates the most severe source of look-ahead bias. If I utilized the contemporaneous year-t network, the open-source release itself could force peer firms to rewrite their product descriptions, meaning the treatment assignment would endogenously reflect the outcome under study. Anchoring the network at $t-1$ forces the treatment assignment to be strictly measurable from information available prior to the event. The Kogan intensity metric is measured over a narrow announcement window and is therefore predetermined relative to the accounting, patenting, and labor outcomes studied here, even though its announcement-window return component is not observable before the release itself. Treatment assignment thus rests on pre-event information, and the intensity weight is fixed before any of the outcomes are realized.

Peers operating within the TNIC network differ systematically from the broader Compustat universe along critical dimensions such as firm age, growth orientation, and R&D intensity. Consequently, an unadjusted comparison to all non-peer firms would conflate the treatment effect with selection bias based on firm characteristics. I therefore construct a matched control group via one-to-one nearest-neighbor propensity score matching estimated at $t-1$. The candidate control pool consists of all Compustat firms outside the focal firm's TNIC network in the pre-event year. Propensity scores are estimated using Market Capitalization, the Market-to-Book Ratio, Leverage, and Total Assets, the covariates whose pre-event balance is reported in Table 2. Matching at $t-1$ ensures that control firms resemble the treated peers based on pre-treatment fundamentals rather than realized outcomes.[9]

The final stacked panel concatenates the cohort-level event panels around each of the approximately 1,700 eligible releases. All continuous financial variables are normalized by total assets or sales and winsorized at the 1st and 99th percentiles to mitigate the influence of outliers.

[8] A small share of treated peers are themselves focal firms in other release events; for example, Oracle and IBM appear as peers in Microsoft's release windows, and vice versa. As a robustness check, I drop the roughly 13% of (focal, peer, event) observations in which the peer firm is itself a focal in the sample. The significantly negative complement-peer CAR is essentially unchanged, from -0.257% to -0.269% over $[0,+3]$, and the Substitute-minus-Complement difference remains positive at $+21$ basis points (versus $+34$ in the full sample); the narrowing comes from the substitute side, part of whose small positive reaction reflects focal firms appearing as substitute peers in other events. The complement-peer compression that drives the asymmetric peer response is therefore not an artifact of cross-event focal-as-peer contamination.

[9] As a robustness check detailed in Appendix Table IA.3, I re-estimate the main specifications using three alternative empirical constructions, specifically the full unmatched sample, Coarsened Exact Matching, and Entropy Balancing, to verify that the core results are preserved. R&D expenditure remains fully robust across all four specifications, Sales is robust across three of the four, and ROA emerges as the most identification-sensitive outcome.

Table 2 reports the pre-event year ($t = -1$) summary statistics for both the treated and matched-control panels. Fundamental firm characteristics, including Market Capitalization, Market-to-Book Ratio, Leverage, and Total Assets, are closely balanced on every covariate entering the propensity score between treated and control group. The difference in group means is at most 0.12 of a pooled standard deviation for each characteristic (market capitalization 0.07, the market-to-book ratio 0.01, leverage 0.06, and total assets 0.12), comfortably within the 0.25 rule of thumb common in the matching literature. Finally, consistent with the earlier validation in Table 1, the distributional statistics for the analytic weight $|\kappa_r|$ confirm a well-behaved, right-skewed, and approximately log-normal shape, which strictly supports its deployment as a continuous scalar weight in the main empirical specifications.

[Insert Table 2 Here]

4.6 Empirical Strategy: Stacked Difference-in-Differences

To identify the causal impact of open-source software releases by focal technology firms, I employ a stacked difference-in-differences approach. This methodology directly addresses the potential biases inherent in standard two-way fixed effects models when treatment timing is staggered and treatment effects are heterogeneous across cohorts (Baker et al., 2022). The final dataset is constructed by concatenating the individual event cohorts into a single master stacked panel. To ensure comparability across heterogeneous firms and to mitigate the influence of extreme outliers, all continuous financial variables are normalized by total assets or total sales and winsorized at the 1st and 99th percentiles.

The primary empirical specification is modeled as follows:

$$Y_{i,c,t} = \alpha + \beta(Treat_{i,c} \times Post_{c,t}) + \Gamma X_{i,t} + \mu_{i,c} + \lambda_{c,t} + \varepsilon_{i,c,t} \tag{1}$$

where $Y_{i,c,t}$ represents the outcome of interest for firm i in cohort c at time t. A cohort c is defined by a specific eligible open-source release event, with a target observation window from $t = -3$ to $t = +3$ years relative to the release year. The stacked panel is unbalanced in event time: each cohort contributes the firm-years available within the 2010–2024 data window, so cohorts released near the end of the sample contribute fewer than three post-release years, and the regression is

estimated on this unbalanced event-time structure rather than on a window forced to be symmetric for every cohort. The interaction term $Treat_{i,c} \times Post_{c,t}$ is the primary variable of interest. $Treat_{i,c}$ is an indicator equal to one if firm i operates in the same TNIC product market as the focal releasing firm in the pre-treatment year, representing the exposed competitor. $Post_{c,t}$ is an indicator taking the value of one for the periods following the open-source release where $t \geq 0$. To account for the heterogeneous economic significance of each release, the observations in this specification are weighted by the absolute value of the cohort-level Kogan intensity, $|\kappa_c|$, as defined in Section 4.1.

The vector $X_{i,t}$ represents a set of time-varying firm-level controls, specifically Market Capitalization, the Market-to-Book ratio, and Leverage, ensuring that the estimates are not driven by shifting financial conditions. I control for unobservable confounders by including cohort-firm fixed effects, $\mu_{i,c}$, which absorb time-invariant firm characteristics within each specific event stack. Furthermore, I include cohort-time fixed effects, $\lambda_{c,t}$, to strictly control for macroeconomic shocks and time-series trends specific to each event window. Standard errors are clustered at the firm level to account for within-firm serial correlation over time.

5 Empirical Results

This section tests the mechanism chain laid out in Hypotheses 3 through 5 of Section 3. Because the specifications in this section pool substitute and complement peers, the estimates recover the average treatment effect over the full peer set. Under Proposition 2, this average is driven by the complement-peer subsample. The first subsection documents the top-line revenue compression predicted by Hypothesis 3. The second subsection shows that peers absorb the revenue shock by cutting direct costs, producing a short-run rise in asset-scaled profitability that completes Hypothesis 3. The third subsection traces the engineering labor reallocation toward the focal firm's open-source repositories predicted by Hypothesis 4. The final subsection documents the retreat in exploratory R&D and the aging of the patent portfolio predicted by Hypothesis 5. Section 6 then partitions the peer set along the Proposition 2 cut and confirms that the average effects documented below load on the complement-peer partition, with the substitute-peer subsample essentially flat.

5.1 The Revenue Shock: Impact on Peer Firm Sales

A strategic open-source release places a free, production-grade substitute into the product market where peer firms sell proprietary alternatives. If this commoditizes the segment as the framework predicts, customers facing a zero-price option of comparable quality would abandon the proprietary version, eroding peer pricing power and compressing peer revenue. I test this prediction on peer firm sales.

Table 3 summarizes how peer firm sales respond to the competitive shock of open-source entry. To ensure the robustness of this revenue impact, I present the results progressively from a baseline OLS model to a saturated specification that addresses both unobserved heterogeneity and the economic heterogeneity of the event. In column (1), using a baseline difference-in-differences specification without fixed effects, I find that normalized sales decrease by 0.049 in the treated group during the post-period compared to the control group, statistically significant at the 1% level.

To rigorously absorb unobservable confounders, I introduce strict Cohort-Firm and Cohort-Time fixed effects and firm-level controls in column (2), with standard errors clustered at the firm level. Under this saturated model, sales decrease by 0.046 in the treated group relative to the control group, statistically significant at the 5% level. In column (3), I further weight each observation by the absolute value of the cohort-level Kogan (2017) intensity, $|\kappa_c|$, so that more economically significant releases receive greater weight in the estimation. The Kogan-weighted coefficient is -0.060, statistically significant at the 1% level. Economically, a 0.060 reduction in sales scaled by total assets is substantial. Relative to the sample mean of 0.728, this translates to an 8.2% decline in the overall revenue generation capacity of incumbent competitors. The Kogan-weighted estimate is the largest of the three specifications, larger than the saturated unweighted estimate, consistent with the compression being concentrated in the most economically consequential releases, though the magnitude does not rise monotonically across the specification ladder. This top-line compression confirms that when dominant platforms release high-quality open-source software, it acts as a direct substitute that cannibalizes the market share and pricing power of proprietary alternatives.

[Insert Table 3 Here]

To examine the parallel-trends assumption underlying this revenue shock, Figure 4 plots the

dynamic treatment effect on sales relative to the base year. In the pre-treatment years, the estimated coefficients are statistically indistinguishable from zero, indicating no divergent pre-existing trends between the treated and control groups prior to the event. Following the open-source release at year zero, there is a sharp and persistent downward trajectory in peer sales. This dynamic pattern is consistent with the revenue compression following the release rather than reflecting prior firm distress, though flat pre-trends alone cannot rule out every other contemporaneous treated-peer shock.

[Insert Figure 4 Here]

5.2 Defensive Restructuring: Cost Efficiency and Profitability

An 8.2 percent structural decline in sales presents a substantial operational challenge to incumbent firms. The economics of open-source software provides a mechanism for these firms to absorb this top-line shock. Theoretical and empirical literature suggests that open-source releases generate substantial technological spillovers, allowing rival firms to free-ride on the focal platform's R&D investments (Nagle, 2019; Emery et al., 2025). By adopting the newly released open-source infrastructure, incumbent firms can replace expensive proprietary software inputs, eliminate redundant development efforts, and achieve significant technological cost efficiencies (Wen et al., 2016; Haese and Peukert, 2025).

Hypothesis 3 predicts that peers absorb the revenue shock not by recovering on the top line but by compressing the cost base, so that asset-scaled profitability rises in the short run even as the competitive position deteriorates. Table 4 tests the two halves of this prediction directly: cost of goods sold scaled by assets, the variable through which the absorption operates, and return on assets, the resulting bottom-line outcome. The prediction concerns asset-scaled profitability and the direct-cost compression that drives it, not conventional operating margin: the test measures whether cost of goods sold falls faster than sales, widening the gross spread, and whether return on assets rises, rather than decomposing operating income into its selling, general, administrative, and other components, which the data do not separately identify here.

Columns (1) and (2) report the cost-channel result. In the baseline column (1), cost of goods and services decreases by 0.045, significant at the 1% level. The saturated column (2) tightens

the estimate to -0.054, also significant at the 1% level, which translates to a 5.4 percentage point reduction in normalized direct costs, or a 16.5% decline relative to the sample mean of 0.328. The free open-source infrastructure substitutes for proprietary inputs that peer firms would otherwise have produced or licensed, and the substitution shows up as an immediate and economically large compression of variable costs.

Columns (3) and (4) trace the bottom-line consequence. The baseline column (3) yields a coefficient of 0.018, significant at the 1% level, and the saturated column (4) yields 0.053, significant at the 5% level, implying a 5.3 percentage point improvement in asset profitability for treated peer firms in the post-period. Because column (3) of Table 3 establishes that the same firms simultaneously lose roughly 8.2% of normalized sales, this profitability gain cannot reflect a top-line recovery. The cost compression in columns (1) and (2) is large enough not only to absorb the revenue shock but to deliver a short-run rise in asset-scaled profitability on top of it, exactly the pattern Hypothesis 3 predicts.

[Insert Table 4 Here]

To validate the sequence of this defensive restructuring, Figure 5 plots the dynamic treatment effects for both ROA and COGS. The evidence supports the parallel trends assumption, as the pre-treatment coefficients for both variables remain flat and statistically indistinguishable from zero prior to the event year. Following the open-source release, Panel A illustrates a persistent downward trajectory in the cost of goods sold over the subsequent three years. Mirroring this timeline, Panel B displays a consistent upward trajectory in ROA. This dynamic symmetry provides evidence that the observed profitability rise is not driven by organic business growth but is the direct consequence of shrinking the firm's cost base via open-source adoption in response to the exogenous revenue shock.

[Insert Figure 5 Here]

This short-run profitability rise admits two readings. A benign interpretation treats the cost compression as a productivity gain that lifts the affected firms, consistent with the open-source-as-spillover literature (Nagle, 2019; Hoffmann et al., 2024). A more guarded interpretation, drawing on the literature on harvested declining segments (Christensen, 1997; Harrigan, 1980), reads the

same pattern as the accounting signature of managed retreat: variable costs fall and proprietary investment is withdrawn, and short-run profitability rises mechanically even as the franchise hollows out. The subsequent subsections adjudicate between these readings: Sections 5.3 through 5.4 test whether the cost compression coincides with reallocation of engineering effort and a withdrawal from frontier innovation, the symptoms of retreat, and Section 6.3 closes the question by asking whether the equity market prices the affected firms as winners or as losers.

5.3 Labor Co-optation: Developer Mobilization Toward the Focal Platform

The cost cutting documented in the previous subsection raises an immediate structural question: if peer firms preserve their product lines while shrinking internal overhead, where does the engineering workforce go? Labor reallocation theories suggest that dominant platforms can co-opt the proprietary talent of their rivals by setting open technological standards (Gupta et al., 2024), since rivals that need to remain compatible with the new industry baseline have an incentive to redirect their developers to the platform's open-source ecosystem (Lerner and Tirole, 2002). This subsection tests that mechanism using real-time developer engagement on GitHub.

Hypothesis 4 predicts that complement peers redirect engineering effort onto the focal platform's open-source repositories to keep their proprietary products compatible with the evolving standard. Table 5 tests this prediction on the full sample of approximately 1,700 cohorts, with observations weighted by the absolute Kogan (2017) innovation value to anchor the estimate on the most economically influential releases. Per-developer activity is split into active structural contributions (pushes, pull requests, issues, issue comments, and forks of the focal repository) and passive engagement (stars and watches), so the active and passive coefficients together provide an internal check: labor reallocation should raise active contribution but not mere attention. Column (4) reports the effect on the count of working developers at peer firms, separating an expanded contributor base from existing contributors working harder.[10]

In column (1), total developer engagement, measured by any work per developer, rises by 1.145 in the treated group during the post-period relative to the control, significant at the 1%

[10]The unweighted variant is reported in Internet Appendix Table IA.4; the unweighted estimates are larger in magnitude and tighter in significance, indicating that the average response across all releases is at least as strong as that on the high-intensity subset. The corresponding full-sample dynamic event-study is reported in Internet Appendix Figure IA1.

level, a 36.7% increase relative to the full-sample mean of 3.122 contributions per developer. The mobilization is driven almost entirely by active structural contributions: in column (2), active contributions (pushes, pull requests, issues, issue comments, and forks of the focal repository) rise by 1.134, significant at the 5% level, accounting for virtually the entire change in total engagement. The active developer base also expands, with the log number of working developers rising by 0.141 in column (4), significant at the 1% level and corresponding to a 15.1% increase in active development staffing on the focal repositories. The passive-engagement coefficient in column (3) is statistically indistinguishable from zero, indicating that the response is concentrated in active rather than passive engagement.

[Insert Table 5 Here]

The dynamic counterpart of this static estimate is reported in Internet Appendix Figure IA1. Peer GitHub activity rises over the sample period in step with the secular expansion of corporate GitHub adoption from 2014 to 2017, and the response is concentrated in active rather than passive contributions, consistent with peers redirecting a meaningful share of their engineering effort onto the focal firm's commons rather than absorbing the revenue shock through cost cuts alone. The next question is whether this reallocation also reshapes the rival's internal pipeline for frontier research.

5.4 The Long-Term Cost: Innovation Retreat

The reallocation of engineering effort onto the focal platform's commons does not merely move developer hours from one accounting bucket to another; it hollows out the rival's capacity for independent discovery. With less proprietary engineering capacity to commit, internal returns to frontier research fall and the rival winds down its in-house R&D pipeline (Cassiman and Veugelers, 2006; Benner and Tushman, 2002). This subsection tests that prediction at the financial margin where it should appear first, peer firms' R&D expenditure scaled by total assets.

Hypothesis 5 predicts that this contraction shows up first as a reduction in peer firms' R&D expenditure, since the marginal return to proprietary *C*-innovation falls once the focal release makes the comparable proprietary output redundant. Table 6 reports the estimate on a three-column ladder, from a baseline OLS difference-in-differences to a saturated specification with Cohort-Firm

and Cohort-Time fixed effects, firm-level controls, and Kogan-intensity weighting. The coefficient on Treated $\times$ Post is -0.006 in column (1), -0.012 in column (2), and -0.014 in column (3), all significant at the 1% level. The Kogan-weighted estimate translates to a 1.4 percentage point reduction in R&D intensity, or a 13.1% decline relative to the sample mean of 0.107. The signal survives intact across all three columns, indicating that the R&D withdrawal is not an artifact of the weighting choice.

[Insert Table 6 Here]

To validate the causal nature of this innovation contraction, Figure 6 plots the dynamic treatment effect on R&D expenditure. The evidence supports the parallel trends assumption, as the pre-treatment coefficients remain flat and statistically indistinguishable from zero prior to the event year. Following the open-source release, the trajectory turns persistently downward over the subsequent three years, with no reversion toward the baseline. This dynamic decline provides evidence that the contraction in research investment is not driven by a transitory accounting adjustment but is the direct consequence of the focal platform's open-source entry.

[Insert Figure 6 Here]

The R&D withdrawal carries a measurable consequence for the stock of patentable innovation that peer firms produce. With less capital committed to frontier research, new exploratory filings slow and the average age of the patent portfolio rises within the three-year post-event window. Internet Appendix Table IA.5 reports estimates for the average age of the overall patent portfolio and the average age of the exploratory portfolio. The Kogan-weighted estimates are 0.323 and 0.792 years respectively, both significant at the 1% level. The aging effect on the exploratory portfolio is more than twice as large as the effect on the portfolio as a whole, consistent with the prediction that the focal release commoditizes precisely the leading edge that peers had previously been racing to patent. The pattern is not a stoppage of patenting altogether but a tilt of the surviving filings toward mature, defensive technology, the standard accounting signature of a firm winding down frontier exploration while preserving existing intellectual property coverage.

6 The Two-Sided Strategy: Rival Cannibalization and Focal-Firm Gain

The baseline estimates document an average compression of rival peer outcomes following focal-firm open-source releases, but the strategic logic of commoditizing one's complement makes two cross-sectional predictions, not one. The first prediction is about who pays: a uniform negative response across all rivals would be equally consistent with generic predation or with a quality signal that updates every observer's prior in the same direction. What separates the commoditize-your-complement strategy from these alternatives is the cross-section of who absorbs the loss: the framework predicts that the harm concentrates on firms whose revenue resides in the segment the release commoditizes, and leaves essentially untouched the firms whose revenue resides in the focal firm's profit center. The second prediction is about who benefits: if the release succeeds in expanding consumption of the commoditized complement, the focal firm's own profit-center sales should rise with the cumulative dose of release, while its other segments respond little or not at all.

This section tests both predictions. I first partition the treated peer cohort along the Substitute–Complement boundary defined in Section 3 and re-estimate the real-side effects within each cell, then close the causal chain on the equity market: if investors price the release as a wealth transfer from complement peers to the focal platform, the announcement-window cumulative abnormal returns should follow the same partition. I close the section by turning to the focal-firm side of the strategy, testing whether the focal firms' headline financial outcomes rise after the release and whether the rise concentrates in the core business segments the release is designed to expand. Together, these analyses translate the average treatment effect into the two-sided wealth transfer the framework predicts: a targeted compression of complement rivals, and a corresponding expansion of the focal firm's core demand.

6.1 Classifying Peers into Substitutes and Complements

The distinction between a substitute and a complement is constructed from two metrics that characterize the economic relationship between each peer firm and the releasing focal firm. The first metric is technology overlap. Following the foundational approach of Jaffe (1986) and subse-

quent literature on technological proximity (Bloom et al., 2013; Bena and Li, 2014), this overlap is the TF-IDF cosine similarity between the peer firm's and the releasing focal firm's Cooperative Patent Classification (CPC) profiles. Each firm's CPC profile is the vector of its patent counts across four-character CPC subclasses, constructed from USPTO patents filed over 2010–2023, linked to firms through the patent-assignee-to-*permno* crosswalk; the counts are TF-IDF weighted so that ubiquitous subclasses do not dominate and L2-normalized before the cosine is taken. The profile is a slow-moving structural characteristic of a firm's technology, measured over a fixed window rather than re-estimated per release, so the overlap between a given focal firm and a given peer is common to all of that firm's releases. This CPC overlap is the first classification dimension: a peer above the cohort median occupies the same technology layer as the releasing firm and is a direct technological rival, while a peer below the median occupies an adjacent layer.

The second metric is the sign of the peer's stock-return co-movement with the focal firm, measured over the pre-event window $[t-250, t-30]$ after netting out industry returns. The use of return co-movement to identify economically linked firm pairs follows a well-established literature on information transmission and production complementarity (Lee et al., 2024). Positive co-movement indicates that the two firms' equity values have historically moved together, consistent with a complementary economic linkage; negative co-movement indicates that they have moved oppositely, consistent with substitution. This pre-event measure proxies the historical economic relationship and is not itself a statement about how investors price the release.

I combine the two metrics and retain only the peers on which they agree. A Substitute is a peer with above-median CPC overlap and negative return co-movement: a firm in the same technology layer whose value moves opposite to the focal firm. A Complement is a peer with below-median CPC overlap and positive return co-movement: a firm in an adjacent layer whose value moves with the focal firm. Peers for which the technology and co-movement signals disagree are set aside as ambiguous and contribute to neither estimate. Requiring both signals to concur sharpens the economic interpretation of each group at the cost of sample size.

This classification directly tests the theoretical prediction. While traditional competitive models suggest the strongest revenue impact should fall on direct substitutes, the commoditization of the complement theory predicts the opposite. The open-source release is designed to erode the pricing power of firms whose products were complementary inputs to the focal platform, convert-

ing paid complements into free alternatives.

6.2 Targeted Impact on Complement Peers

The framework predicts that the real-side effects established in Section 5 load asymmetrically on the two peer groups. Complement peers, whose proprietary revenue resides in the segment the focal release commoditizes, should bear the bulk of the revenue compression and the cost-base adjustment. Substitute peers, whose revenue resides in the focal firm's protected profit center, should show little or no compression on the directly commoditized segment, with at most a weaker cost-side adjustment from adopting the same free infrastructure. I test this prediction on the two headline financial channels established in the pooled baseline, Sales, through which pricing power on the directly affected segment first compresses, and Cost of Goods Sold, through which the free focal infrastructure substitutes for proprietary inputs. These two outcomes are the cleanest test of the partition because both are the financial margins on which the commoditization should leave its first measurable trace.

Table 7 reports the decomposition for Sales and COGS. The specification is identical to the pooled baseline, estimated separately within each subsample. Columns (1) and (2) report the Sales results. For Substitute peers, the treatment coefficient is -0.055, statistically indistinguishable from zero. For Complement peers, the coefficient is -0.121, significant at the 1% level. The magnitude of the decline is roughly 2.2 times larger for Complement peers. This pattern confirms the partition prediction: the revenue compression concentrates on the peers that occupy an adjacent technology layer and whose pre-event returns moved with the focal firm, the group the classifier identifies as complements.

Columns (3) and (4) report the decomposition for COGS. Both groups cut direct costs after the release, but the Complement peer cut of -0.077 is substantially larger than the Substitute peer cut of -0.047. The asymmetry is economically intuitive. The free infrastructure released by the focal firm is a generalized input that any technologically adjacent peer can substitute into its cost base, which is why both groups show a cost reduction. Only Complement peers, however, lose the paid pricing surface that the release directly commoditizes, which is why the magnitude is larger on their side.

[Insert Table 7 Here]

The substitute–complement contrast does not hinge on the particular signals used to draw the boundary. As a robustness check on the classification itself, Internet Appendix Table IA.9 re-estimates the decomposition under an alternative partition in which each peer is labeled by a panel of three large language models that read its core business rather than its patent-class overlap and return co-movement. The complement-peer compression is, if anything, sharper under this fundamentals-based labeling, the substitute-peer effects remain statistically indistinguishable from zero across sales, costs, profitability, and R&D, and the short-run profitability rise that is muted under the mechanical partition becomes large and significant for complements. Because a relabeling uncorrelated with the truth could only attenuate the contrast, the fact that an independent economic classification sharpens it indicates that the asymmetry reflects genuine economic complementarity rather than an artifact of the mechanical proxy.

The dynamic counterpart of these estimates is shown in Figure 7. Panel A plots the event-study trajectory for Complement peers, and Panel B plots it for Substitutes. The Complement peer sales path is flat prior to the release, drops at year zero, and remains persistently below the pre-period base. The Substitute peer path shows no meaningful deviation from zero. This graphical evidence localizes the full weight of the sales effect documented in the baseline results to the Complement partition.

[Insert Figure 7 Here]

6.3 Market Reaction and Wealth Transfer

The real-side results establish that the revenue and cost restructuring fall overwhelmingly on complement peers. If investors correctly understand the economics of commoditizing the complement, they should reprice the affected firms along the same partition in real time. This subsection tests that prediction by examining the cumulative abnormal returns (CAR) of focal firms, substitute peers, and complement peers around each release, and by translating these percentage returns into aggregate dollar-value impacts. Together, these estimates complete the causal chain by showing that the market's immediate pricing of the shock mirrors the product-market partition shown above.

One refinement applies to this test. The asset-price partition retains the technology-overlap dimension of the classification, the above-median CPC split together with a same-sector restriction on complements, but excludes the pre-event return-co-movement criterion used in the product-market partition of Section 6.1. Classifying peers on pre-event returns and then measuring their abnormal returns would mechanically induce the very pattern under study, so dropping that criterion keeps the partition independent of the return outcome.

I report the test on the full sample of approximately 1,700 open-source releases. Two methodological choices are reflected in Table 8. First, I use value-weighted CARs to reflect the economic weight of each firm, since the outcome of interest, aggregate wealth destruction in the complement pool, is intrinsically dollar-weighted; an equal-weighted version would conflate the large-cap incumbents that bear most of the economic impact with a long tail of small peers whose dollar stakes are negligible. Second, the standard practice of combining pre-event and post-event days into a single CAR can conceal a sharper pattern. I therefore report a within-window placebo over $[-1,0)$ alongside three cumulative post-event windows $[0,+k]$ for $k \in \{1,3,5\}$, allowing the pre-event flatness and the post-event response to be read off the same table directly.

[Insert Table 8 Here]

The pre-event window $[-1,0)$ is statistically indistinguishable from zero for all three groups, a short pre-event placebo consistent with no information leakage in the trading day before the announcement. Over the headline post-event window $[0,+3]$, the three groups separate along the partition the framework predicts. Focal firms gain 0.32 percent on average, significant at the 1% level. Complement peers lose 0.26 percent, also significant at the 1% level. Substitute peers gain 0.09 percent, smaller in magnitude but also significant at the 1% level, consistent with the framework's prediction that a complement-market expansion raises demand for the focal segment in which substitute peers also operate. The pattern is robust across the two alternative cumulative windows reported in the same table: the complement loss is largest at $[0,+3]$, smaller in absolute terms at $[0,+1]$, and partially reverts by $[0,+5]$ as part of the second downward leg is absorbed.

[Insert Figure 8 Here]

Figure 8 visualizes the same decomposition as the cumulative average abnormal return from $t=-1$ through $t=+5$ for each of the three groups. The pre-event day $t=-1$ ties all three

trajectories to a common baseline near zero. The focal path rises monotonically toward its peak near $t = +3$, the substitute path traces a shallow positive drift, and the complement path drops at $t = 0$, partially reverts at $t = +1$ and $t = +2$, and reaches its deepest trough at $t = +3$ before recovering. The visual separation around the release reproduces the same partition that organizes the real-side outcomes in Section 6.2, with the complement peers absorbing the bulk of the price reaction. The long-window robustness counterpart of this figure, restricted to the top 100 strategic releases and extended through day $+20$, is reported in Internet Appendix Figure IA2.

[Insert Figure 9 Here]

Figure 9 translates the percentage returns into aggregate dollar-value impacts following the wealth-transfer accounting in Emery et al. (2025) and Hoffmann et al. (2024). To avoid double-counting the same large platforms that recur across many release cohorts, I aggregate per unique firm: within each partition I take each firm's average CAR over the window, multiply by its pre-event market capitalization, and sum across firms. Over the headline window $[0, +3]$, the releasing platforms add an aggregate \$42 billion in market value, complement peers lose an aggregate \$4 billion, and substitute peers are economically small and second-order. The repricing is therefore asymmetric in the same direction as the product market: the platform layer captures value and the complement layer loses it, while the substitute layer, which competes in the focal firm's protected segment rather than the commoditized one, is left largely untouched.

Taken together, the market-reaction evidence completes the mechanism shown in the preceding subsections. The same complement firms that suffer the largest sales compression and execute the sharpest cost restructuring are the firms whose equity value the market marks down in real time. Substitute peers, whose real-side outcomes are economically muted, receive a small positive market reaction consistent with the spillover term in Proposition 2. This alignment between the product-market partition and the contemporaneous market-implied valuation shock constitutes the paper's strongest evidence that open-source releases operate as targeted commoditization events rather than as undifferentiated technological spillovers.

6.4 Do Focal Firms Benefit from Releasing Open-Source Code?

The headline results so far document the rival side of the strategy: complement peers absorb the bulk of the compression. The complementary question is whether the focal firm itself benefits from these strategically valuable releases, and through what channel. I close this section with two analyses on the focal-firm side: first whether the focal firms' aggregate financial performance rises after the release, and then whether the rise concentrates in the business segments the release is designed to expand.

The focal-firm set is determined by data coverage rather than by an ex-ante restriction. Among large-cap technology firms identified as candidate strategic releasers, I retain those with at least one open-source release in the sample window and a pre-treatment window of at least three fiscal years in Compustat. One firm (Atlassian, with first eligible release in 2024) is excluded because its post-treatment window is empty within the sample period, leaving 15 focal firms: Adobe, Alphabet, Amazon, Apple, Cisco, IBM, Meta, Microsoft, Netflix, Nvidia, Oracle, Qualcomm, Salesforce, Tesla, and Twilio. Other candidate firms either ran no open-source release in the sample window or completed their initial public offering too recently to satisfy the pre-treatment-window requirement.

The 15 focal firms have a different panel structure from the peer-side stacked design. Each focal firm enters as a single treated unit with a single first-release date, so the stacked-cohort framework of Section 4.6 does not apply. I instead estimate a standard staggered Difference-in-Differences on an annual panel of the 15 focal firms together with a propensity-score-matched control pool drawn from the Compustat tech-firm pool (SIC 73, 35, 36, 38), excluding any firm that itself completed a open-source release in the sample window.[11] The matching yields 53 unique controls; the panel spans fiscal years 2010 through 2023. Each focal firm's first-treatment year is its first eligible open-source release year: 2015 for Adobe, Alphabet, Amazon, Apple, Cisco, IBM, Meta, Microsoft, Netflix, Nvidia, and Oracle; 2017 for Salesforce; 2018 for Tesla; and 2020 for Qualcomm and Twilio.

The specification is

[11] For each focal firm, I select 5 nearest-neighbor controls on pre-release ln(total assets), ln(sales), ROA, market-to-book, and leverage, with exact two-digit SIC matching, a 0.25-standard-deviation caliper, and matching with replacement.

$$Y_{i,t} = \alpha + \beta_F \left(Focal_i \times Post_{i,t} \right) + \mu_i + \lambda_t + \varepsilon_{i,t} \tag{2}$$

where $Y_{i,t}$ is the Log Sales or Return on Assets of firm i in fiscal year t, $Focal_i$ equals one for the 15 focal firms and zero for matched controls, and $Post_{i,t}$ equals one for fiscal years on or after firm i's first-treatment year. μ_i is a firm fixed effect and λ_t is a fiscal-year fixed effect. Standard errors are clustered at the firm level. Because $Focal_i = 0$ for every matched control, the interaction $Focal_i \times Post_{i,t}$ is zero for controls throughout: the controls carry no own treatment date and enter only through the firm and fiscal-year fixed effects as the untreated comparison, while $Post_{i,t}$ switches on only for the 15 treated focal firms at their respective first-treatment years. The coefficient β_F therefore identifies the average post-release shift in the focal-firm outcome relative to matched controls, after absorbing firm-specific levels and common annual shocks.

A standard concern with this two-way fixed-effect estimator is that it can be biased when treatment timing is staggered and treatment effects are heterogeneous across cohorts, because already-treated units serve as implicit controls for later-treated cohorts (Goodman-Bacon, 2021; Sun and Abraham, 2021). I address this by re-estimating the same specification with the Callaway and Sant'Anna (2021) doubly-robust ATT estimator (CS-DR), which restricts the comparison group to the 53 never-treated matched controls and weights only clean 2×2 cohort-time differences, ruling out the forbidden comparisons that drive the TWFE bias.

Table 9 reports both estimators side by side. Under TWFE, focal-firm Log Sales in column (1) is 0.838 higher in the post-release period than for matched controls, statistically significant at the 1% level, and Return on Assets in column (3) is 0.042 higher, statistically significant at the 10% level. The doubly-robust estimates in columns (2) and (4) confirm the headline result: Log Sales is 0.519 higher and Return on Assets is 0.043 higher, both statistically significant at the 1% level. These magnitudes are economically substantial. Relative to their matched controls, the focal firms expand revenue by roughly 68% over the post-release window and improve return on assets by an average of 4.3 percentage points.

Taken together, the two headline outcomes show that the commoditize-your-complement strategy delivers both a top-line and a bottom-line gain to the focal platform. The release expands revenue relative to matched controls, and the expansion is large enough to lift return on assets at

the same time rather than dissipating into cost. The same firms that engineer the rival-side compression documented in the preceding subsections also reap a meaningful financial payoff, confirming that the strategy is not merely defensive but actively value-creating for the focal platform itself.

[Insert Table 9 Here]

Figure 10 reports the dynamic event-study counterpart estimated under CS-DR. The pre-treatment coefficients for both outcomes lie within their confidence bands around zero, consistent with the parallel-trends assumption. The post-treatment trajectory is positive and monotonically increasing for both outcomes through year five after the release, with Return on Assets becoming statistically significant from event time $+1$ onward.

[Insert Figure 10 Here]

The headline financials document that focal firms benefit but do not identify the channel. The commoditize-your-complement framework predicts a specific mechanism: by giving away the layer that previously slowed adoption of the focal firm's profit-center product, the release should expand demand for that product, raising sales in the focal firm's core business segment faster than in its non-core segments. I test this prediction directly with Compustat Segment data over fiscal years 2010 through 2023. Each focal firm reports its business segments in 10-K disclosures, and I hand-classify each segment as Core if it is the firm's profit-center segment, segment A in the framework of Section 3, whose demand the release is designed to expand by commoditizing the complementary segment; for these platform firms the profit center is typically the cloud, platform, or core product franchise. All other segments are Non-Core. The full mapping is reported in Internet Appendix Table IA.6. The test compares the post-release sales path of a focal firm's Core segments against its own Non-Core segments in the same fiscal year, so the comparison group is internal to each firm and no external control firms enter the estimate.

The specification regresses the natural logarithm of segment sales on the interaction of the segment's Core indicator with a post-release indicator:

$$\ln(Sales_{i,s,t}) = \alpha + \beta\,(IsCore_{i,s} \times Post_{i,t}) + \mu_{i,t} + \nu_{i,s} + \varepsilon_{i,s,t} \tag{3}$$

where i indexes firm, s indexes business segment, and t indexes fiscal year. $IsCore_{i,s}$ equals one for segments classified as Core and zero otherwise, and $Post_{i,t}$ equals one in the fiscal years on or after firm i's first eligible open-source release. Firm$\times$year fixed effects $\mu_{i,t}$ absorb every firm-level confound, including the post-2015 cloud boom, macroeconomic shocks, and firm-specific managerial regimes; firm$\times$segment fixed effects $\nu_{i,s}$ absorb each segment's baseline level. Identification of β therefore comes purely from within-firm-year differential growth of Core relative to Non-Core segments around the release. Standard errors are clustered at the firm$\times$segment level.

Table 10 reports the estimate across two sample windows and two inference choices. Over the full sample in columns (1) and (2), the contemporaneous effect is $\beta = 0.350$, which implies that a focal firm's Core-segment sales rise about 42% more than its Non-Core-segment sales following the release. The estimate is significant at the 1% level whether standard errors are left unadjusted or clustered at the firm$\times$segment level. Restricting the sample to fiscal years 2014 through 2023, the period during which strategic open-source release became a first-order corporate activity at the focal firms, raises the estimate to 0.471 in columns (3) and (4), equivalent to a Core-versus-Non-Core sales gap of roughly 60% and again significant at the 1% level under both inference choices.

[Insert Table 10 Here]

Figure 11 plots the event-study counterpart: the Core-versus-Non-Core differential is flat and statistically indistinguishable from zero in every pre-release year, then rises monotonically from the release onward, confirming that the demand reallocation follows the release rather than preceding it. These two analyses together establish the focal-firm side of the commoditize-your-complement strategy. The aggregate Return on Assets and Sales of the focal firms rise after the release, and within firm and year the core business segments grow faster than the non-core segments the release does not target. The segment design identifies this Core-versus-Non-Core differential rather than separate absolute changes in either group.

[Insert Figure 11 Here]

6.5 Discussion and Robustness

To provide concrete economic context for this partition, Appendix Table IA.7 presents illustrative examples of firm pairs categorized into the Substitute and Complement buckets. A qualitative review of these matched pairs confirms that the algorithmic classification accurately aligns with the structural realities of the product market. Firms classified as Complements typically provide adjacent or interdependent services to the focal platform, whereas those classified as Substitutes offer direct, competing alternatives.

A potential concern with the two-dimensional partition is that setting aside the ambiguous off-diagonal peers, which removes roughly two-thirds of the classifiable peers, might itself manufacture the asymmetry. To address this, Appendix Table IA.8 re-estimates the decomposition under a simpler one-dimensional classification that uses CPC technology overlap alone and retains every peer. The Sales and COGS results are unchanged in direction and significance: the complement-peer revenue compression remains large and significant while the substitute-peer effect stays small and insignificant, confirming that the concentration of the shock among complement firms is not an artifact of dropping the ambiguous peers.

This decomposition refines the paper's central empirical finding. For the two financial channels decomposed here, Sales and COGS, the average treatment effects documented in the baseline analysis do not represent a diffuse industry spillover but a targeted compression of peer firms whose products operate in a complementary relationship to the releasing platform; under Proposition 2 the remaining channels, which are estimated on the pooled peer set, are expected to load on the same complement pool. On these decomposed margins the results also rule out conventional head-to-head rivalry as the primary channel of cannibalization, since substitute peers absorb a substantially smaller shock than complement peers do. Ultimately, these focal platforms successfully commoditize the paid complements upon which their own ecosystems previously relied.

7 Conclusion

The modern digital economy is increasingly characterized by a structural shift wherein dominant technology firms invest substantial capital to develop advanced software, only to release it to the public at no cost. While the open-source movement is frequently viewed as a collaborative

public good, this study provides comprehensive empirical evidence of a more calculated economic reality. By integrating the economics of two-sided network effects with the resource-based view of the firm, this paper demonstrates that dominant platforms strategically deploy open-source software to commoditize complementary markets, displacing the proprietary resource advantages of incumbent firms.

The empirical anatomy of this commoditization shock is sequential and structurally significant. Confronted with a structural top-line revenue compression, incumbent peers are forced into an operational restructuring. They reduce direct costs and organizational overhead by substituting proprietary inputs with the newly available open-source infrastructure. While this defensive pivot mechanically inflates short-term profitability, it is not a costless adjustment. Instead, it is mirrored in a substantial reallocation of engineering human capital, as competitor engineers are effectively co-opted and mobilize on platforms such as GitHub to maintain and expand the focal firm's open-source ecosystem.

By outsourcing their core technological foundations to the dominant platform in this way, these rivals systematically hollow out their independent innovation capabilities. Research and development budgets contract, patent portfolios age rapidly, and exploratory innovation at the technological frontier declines.

The heterogeneity analysis confirms that these dynamics are not the byproduct of a diffuse technological spillover. The impact of open-source releases is highly targeted. When partitioning the treated cohort into substitute and complement peers, the revenue compression and cost restructuring are concentrated among firms operating as complements to the focal platform, while direct substitutes remain largely insulated on sales; the labor and innovation channels are estimated on the pooled peer set, which under the framework loads predominantly on the same complement pool. This targeted partition is immediately validated by the financial markets: over the $[0, +3]$ window investors mark up the releasing platforms and mark down complement peers in real time, with the platforms adding roughly \$42 billion in aggregate market value while complement peers lose value and substitutes remain close to unchanged. These asset-price estimates close the causal chain by showing that the product-market partition and the contemporaneous market-implied valuation shock line up on the same firms, confirming the central theoretical prediction that open-source software is strategically targeted to erode the pricing power of complementary layers rather than

head-to-head substitutes.

The transfer has a second arm on the focal firm's own books. The release that compresses complement rivals simultaneously expands the releasing platform's core business: within firm and year, core-segment sales rise about 42 percent more than non-core-segment sales after the release, and firm-level sales and profitability rise relative to matched controls. The market-implied gain to the focal firm is therefore not a windfall but the priced counterpart of a real demand reallocation, channeling the rents freed from the commoditized complement into the platform's profit center.

The historical precedent for this dynamic clarifies what is at stake. Gurley (2026) notes that Amazon Web Services captured the cloud only because open-source commoditization of the operating-system and middleware stack had already eliminated the supplier power that would have allowed IBM, Hewlett-Packard, or Intel to extract rents. In that case the commoditization was exogenous, driven by a developer community rather than by Amazon. The contemporary version studied here is consistent with an endogenous strategy: dominant incumbents release the complementary asset themselves to occupy the operator position. The evidence does not directly observe managerial intent or the internal financing of the release, but the asymmetric repricing and the real-side adjustments line up with this interpretation. The redistributive consequences are the same, but the strategic agency appears to have changed, and the regulatory framework has not caught up.

These findings carry significant implications for both academic theory and antitrust policy. In the strategic management literature, this paper illustrates how cross-market subsidies can be utilized to convert a rival's proprietary assets into obsolete legacy systems. Furthermore, it highlights a critical blind spot in modern antitrust enforcement. Traditional regulatory frameworks focus predominantly on consumer harm via monopoly pricing and output restrictions. However, in the digital platform ecosystem, market power is frequently exercised through the free distribution of complements. By giving away advanced software, dominant platforms avoid traditional antitrust scrutiny while successfully diminishing rivals' revenue, stifling independent exploratory innovation, and absorbing the industry's engineering human capital. Ultimately, this study demonstrates that through the strategic commoditization of the complement, dominant platforms systematically neutralize competitive threats and secure their structural advantage in the digital economy.

References

Adner, Ron, and Rahul Kapoor, 2010, Value creation in innovation ecosystems: How the structure of technological interdependence affects firm performance in new technology generations, *Strategic Management Journal* 31, 306–333.

Aghion, Philippe, Nick Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt, 2005, Competition and innovation: An inverted-u relationship, *The Quarterly Journal of Economics* 120, 701–728.

Alexy, Oliver, Gerard George, and Ammon J. Salter, 2013, Cui bono? the selective revealing of knowledge and its implications for innovative activity, *Academy of Management Review* 38, 270–291.

Alexy, Oliver, Joel West, Helge Klapper, and Markus Reitzig, 2018, Surrendering control to gain advantage: Reconciling openness and the resource-based view of the firm, *Strategic Management Journal* 39, 1704–1727.

Armstrong, Mark, 2006, Competition in two-sided markets, *RAND Journal of Economics* 37, 668–691.

Baker, Andrew C, David F Larcker, and Charles CY Wang, 2022, How much should we trust staggered difference-in-differences estimates?, *Journal of Financial Economics* 144, 370–395.

Baker, Scott R, Brian Baugh, and Marco Sammon, 2023, Customer churn and intangible capital, *Journal of Political Economy Macroeconomics* 1, 447–505.

Balsmeier, Benjamin, Lee Fleming, and Gustavo Manso, 2017, Independent boards and innovation, *Journal of Financial Economics* 123, 536–557.

Barney, Jay, 1991, Firm resources and sustained competitive advantage, *Journal of Management* 17, 99–120.

Bena, Jan, and Kai Li, 2014, Corporate innovations and mergers and acquisitions, *The Journal of Finance* 69, 1923–1960.

Benner, Mary J, and Michael Tushman, 2002, Exploratory and exploitative innovation, learning, and the impact of corporate quality management, *Academy of Management Journal* 45, 676–698.

Bian, B., Q. Huang, Y. Li, and H. Tang, 2025, Data as a networked asset, *Working Paper* .

Bloom, Nicholas, Mark Schankerman, and John Van Reenen, 2013, Identifying technology spillovers and product market rivalry, *Econometrica* 81, 1347–1393.

Boudreau, Kevin, 2010, Open platform strategies and innovation: Granting access vs. devolving control, *Management Science* 56, 1849–1872.

Callaway, Brantly, and Pedro HC Sant'Anna, 2021, Difference-in-differences with multiple time periods, *Journal of Econometrics* 225, 200–230.

Casadesus-Masanell, Ramon, and Pankaj Ghemawat, 2006, Dynamic mixed duopoly: A model motivated by Linux vs. Windows, *Management Science* 52, 1072–1084.

Casadesus-Masanell, Ramon, and Gaston Llanes, 2011, Mixed source, *Management Science* 57, 1212–1230.

Cassiman, Bruno, and Reinhilde Veugelers, 2006, In-house r&d and external knowledge acquisition: complementary or substitute?, *Research Policy* 35, 68–82.

Cennamo, Carmelo, Hakan Ozalp, and Tobias Kretschmer, 2018, Platform architecture and quality trade-offs of multihoming complements, *Information Systems Research* 29, 461–478.

Chesbrough, Henry, and A. K. Crowther, 2006, Beyond high tech: early adopters of open innovation in other industries, *R&D Management* 36, 229–236.

Christensen, Clayton M., 1997, *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail* (Harvard Business School Press, Boston, MA).

Coleman, J., K. Fronk, and K. Valentine, 2025, Corporate adoption of an open innovation strategy: Evidence from github, *Working Paper* .

Conti, Annamaria, Christian Peukert, and Maria Roche, 2025, Beefing it up for your investor? engagement with open source communities, innovation, and startup funding: Evidence from github, *Organization Science* 36, 1551–1573.

Cournot, Antoine Augustin, 1838, *Recherches sur les Principes Mathématiques de la Théorie des Richesses* (Hachette, Paris).

Economides, Nicholas, and Evangelos Katsamakas, 2006, Two-sided competition of proprietary vs. open source technology platforms and the implications for the software industry, *Management Science* 52, 1057–1071.

Eisenmann, Thomas, Geoffrey Parker, and Marshall Van Alstyne, 2011, Platform envelopment, *Strategic Management Journal* 32, 1270–1285.

Emery, Logan P., Chan Lim, and Shiwei Ye, 2025, The private value of open-source innovation, Working paper, SSRN.

Fama, Eugene F, and Kenneth R French, 1992, The cross-section of expected stock returns, *The Journal of Finance* 47, 427–465.

Farrell, Joseph, and Garth Saloner, 1985, Standardization, compatibility, and innovation, *RAND Journal of Economics* 16, 70–83.

Fosfuri, Andrea, Marco S. Giarratana, and Alessandra Luzzi, 2008, The penguin has entered the building: The commercialization of open source software products, *Organization Science* 19, 292–305.

Frank, Murray Z, and Vidhan K Goyal, 2009, Capital structure decisions: which factors are reliably important?, *Financial Management* 38, 1–37.

Gawer, Annabelle, and Rebecca Henderson, 2007, Platform owner entry and innovation in complementary markets: Evidence from intel, *Journal of Economics & Management Strategy* 16, 1–34.

Goodman-Bacon, Andrew, 2021, Difference-in-differences with variation in treatment timing, *Journal of Econometrics* 225, 254–277.

Gupta, A., N. Nishesh, and E. Simintzi, 2024, Big data and bigger firms: A labor market channel, *SSRN Electronic Journal* .

Gurley, Bill, 2026, From Open Source Software to Open Source Strategy: How the Smartest Executives Are Using Open Source Techniques to Optimize Corporate Strategy, `https://p3institute.substack.com/p/from-open-source-software-to-open`, Substack post, May 10, 2026.

Haese, M., and C. Peukert, 2025, Open at the core: Moving from proprietary technology to building a product on open source software, *Working Paper* .

Harrigan, Kathryn Rudie, 1980, *Strategies for Declining Businesses* (Lexington Books, Lexington, MA).

Heath, D., N. Seegert, R. Wuebker, and J. Yang, 2023, Teams and the homophily trap: Evidence from open source software, *Working Paper* .

Henkel, Joachim, 2006, Selective revealing in open innovation processes: The case of embedded Linux, *Research Policy* 35, 953–969.

Hoberg, Gerard, and Gordon Phillips, 2016, Text-based network industries and endogenous product differentiation, *Journal of Political Economy* 124, 1423–1465.

Hoberg, Gerard, Gordon Phillips, and Nagpurnanand Prabhala, 2014, Product market fluidities, technology spillovers, and product market rivalry, *The Review of Financial Studies* 27, 333–371.

Hoffmann, Manuel, Frank Nagle, and Yanuo Zhou, 2024, The value of open source software, Working Paper 24-038, Harvard Business School.

Jaffe, Adam B, 1986, Technological opportunity and spillovers of r&d: Evidence from firms' patents, profits, and market value, *The American Economic Review* 76, 984–1001.

Katz, Michael L., and Carl Shapiro, 1985, Network externalities, competition, and compatibility, *American Economic Review* 75, 424–440.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman, 2017, Technological innovation, resource allocation, and growth, *The Quarterly Journal of Economics* 132, 665–712.

Lee, C. M. C., T. T. Shi, S. T. Sun, and R. Zhang, 2024, Production complementarity and information transmission across industries, *Journal of Financial Economics* 155, 103812.

Lerner, Josh, and Jean Tirole, 2002, Some simple economics of open source, *The Journal of Industrial Economics* 50, 197–234.

Lin, Y.-K., and L. M. Maruping, 2022, Open source collaboration in digital entrepreneurship, *Organization Science* 33, 212–230.

Makaew, Tanakorn, and Vojislav Maksimovic, 2020, Competition and operating volatilities around the world, *Journal of Financial and Quantitative Analysis* 55, 517–547.

Manso, Gustavo, 2011, Motivating innovation, *The Journal of Finance* 66, 2113–2153.

Mkrtchyan, A., J. Bai, R. Dai, and C. Wan, 2025, Creativity without walls: The case of open innovation, *SSRN Electronic Journal* .

Nagaraj, Abhishek, and Henning Piezunka, 2024, The divergent effects of firm sponsorship on open collaboration, *Strategy Science* 9, 277–296.

Nagle, Frank, 2019, Open source software and firm productivity, *Management Science* 65, 1191–1215.

Parker, Geoffrey G, and Marshall W Van Alstyne, 2005, Two-sided network effects: A theory of information product design, *Management Science* 51, 1494–1504.

Roberts, Michael R, and Toni M Whited, 2013, Endogeneity in empirical corporate finance, in *Handbook of the Economics of Finance*, volume 2, 493–572 (Elsevier).

Rochet, Jean-Charles, and Jean Tirole, 2003, Platform competition in two-sided markets, *Journal of the European Economic Association* 1, 990–1029.

Seamans, Robert, and Feng Zhu, 2014, Responses to entry in multi-sided markets: The impact of Craigslist on local newspapers, *Management Science* 60, 476–493.

Shapiro, Carl, and Hal R. Varian, 1999, *Information Rules: A Strategic Guide to the Network Economy* (Harvard Business School Press, Boston, MA).

Spolsky, Joel, 2002, Strategy letter V, Essay, Joel on Software, `https://www.joelonsoftware.com/2002/06/12/strategy-letter-v/`.

Sun, Liyang, and Sarah Abraham, 2021, Estimating dynamic treatment effects in event studies with heterogeneous treatment effects, *Journal of Econometrics* 225, 175–199.

Varian, Hal R., 2004, Economics of information technology, in *The Economics of Information Technology: An Introduction* (Cambridge University Press, Cambridge, UK), Raffaele Mattioli Lectures.

Wen, Wen, Marco Ceccagnoli, and Chris Forman, 2016, Opening up intellectual property strategy: Implications for open source software entry by start-up firms, *Management Science* 62, 2668–2691.

Wen, Wen, and Feng Zhu, 2019, Threat of platform-owner entry and complementor responses: Evidence from the mobile app market, *Strategic Management Journal* 40, 1336–1367.

Zhao, Yue, 2025, Open-source generative ai and firm value: Evidence from the release of deepseek, *Working Paper, Singapore Management University* .

Zhu, Kevin Xiaoguo, and Zach Zhizhong Zhou, 2012, Research note—lock-in strategy in software competition: Open-source software vs. proprietary software, *Information Systems Research* 23, 536–545.

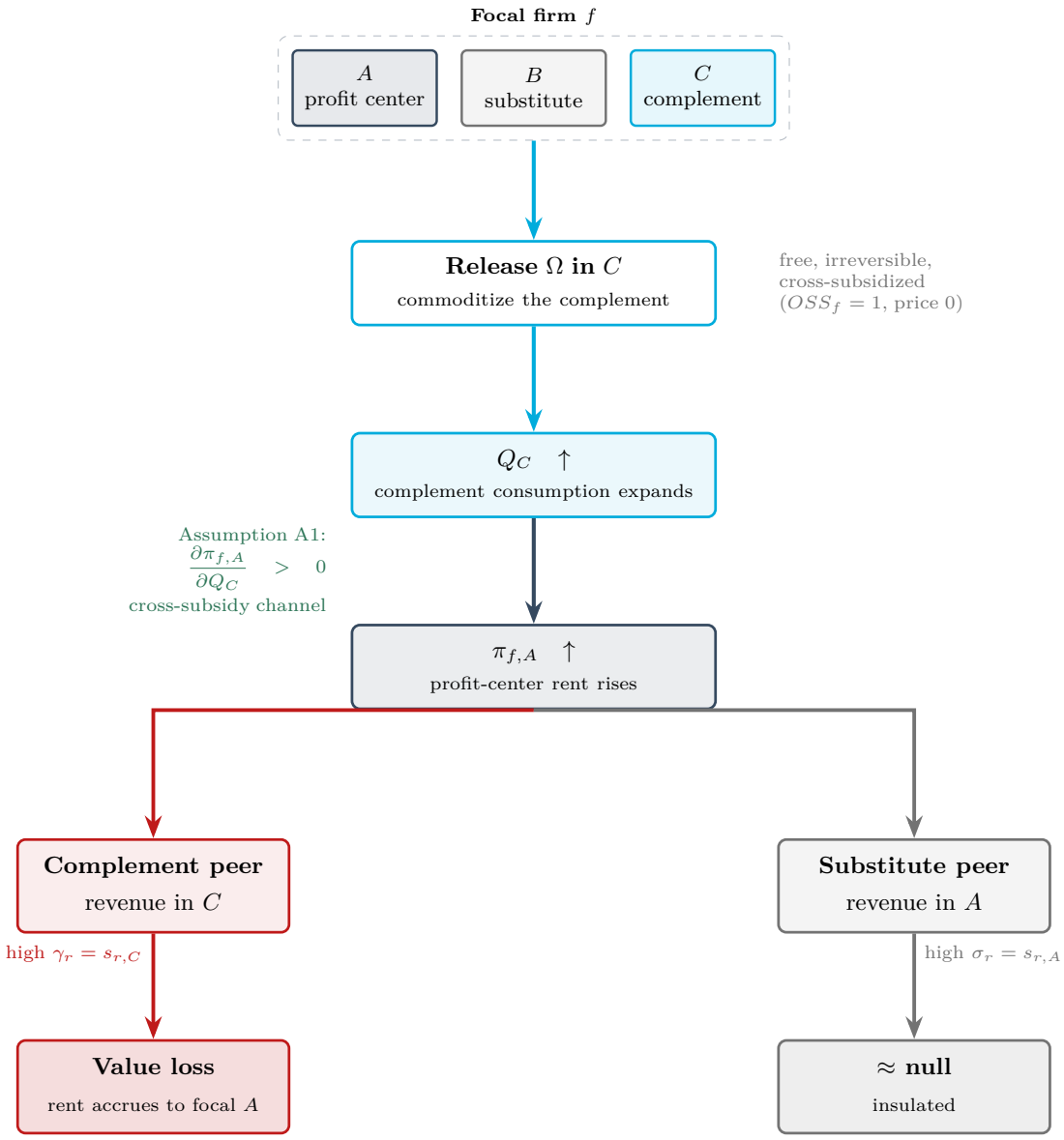


Figure 1. Theoretical Framework: The Commoditize-Your-Complement Mechanism
The focal firm operates across three segments: a profit center A, a substitute market B, and a complement market C. Releasing a free open-source asset Ω in C expands consumption Q_C, which raises the marginal rent of the profit center A through Assumption A1 ($\partial \pi_{f,A}/\partial Q_C > 0$); the release is optimal when this cross-subsidy, together with the standard-setting and talent channels, exceeds the direct loss of C-segment rent (Proposition 1). The same shock has opposite effects on product-market peers depending on where their revenue lives: for a complement peer (high γ_r), Ω undercuts its C-segment price toward marginal cost and compresses its value, which accrues to the focal firm's profit center; for a substitute peer (high σ_r), the only effect is a weak, capped A-side gain (Proposition 2).

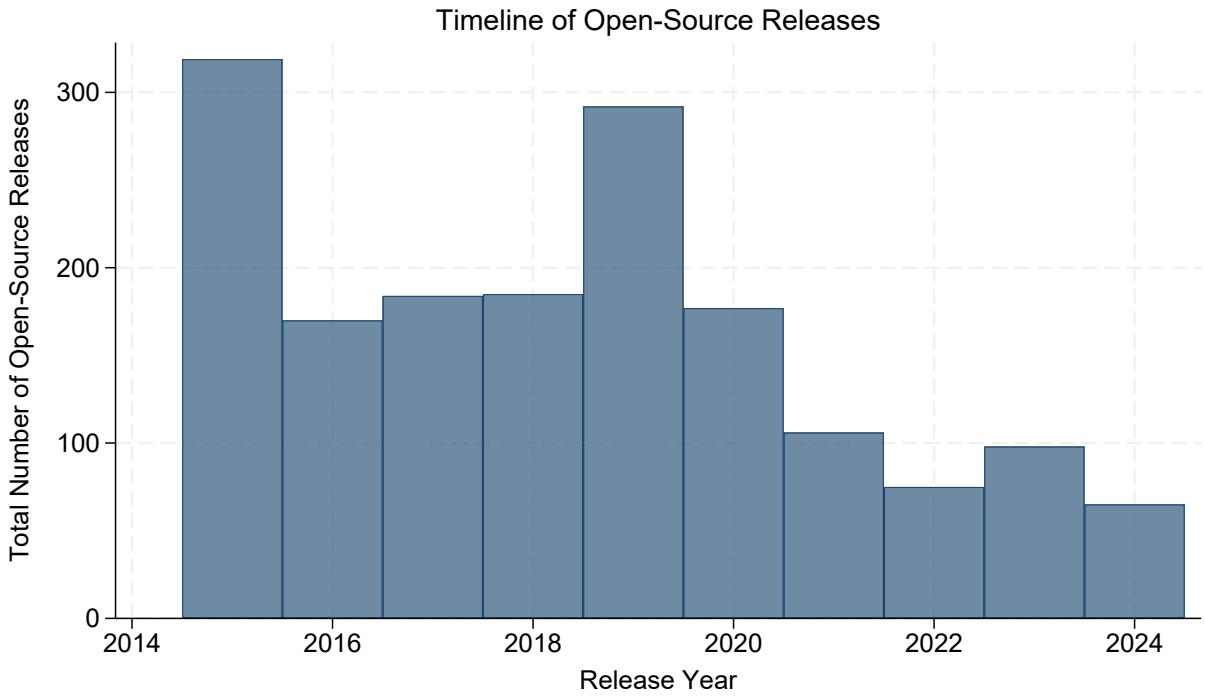
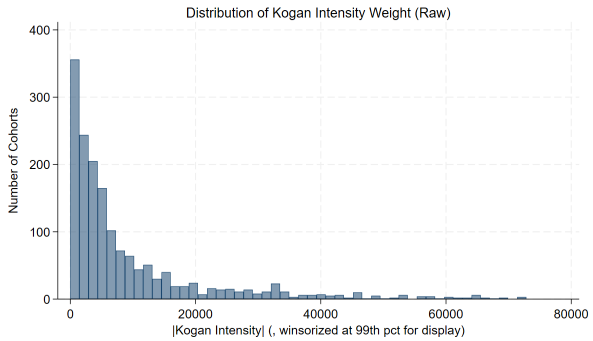
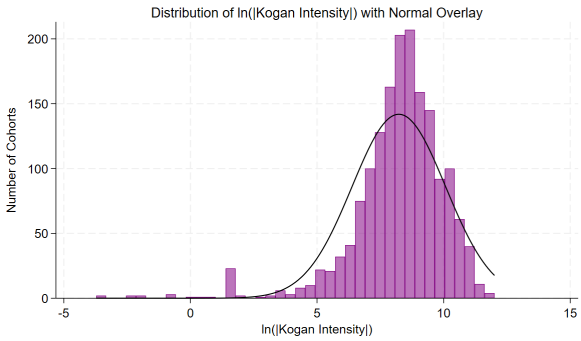


Figure 2. Timeline of Total Open-Source Releases

This figure presents the annual distribution of the approximately 1,700 open-source repositories in the full eligible cohort set released between 2015 and 2024. The histogram reports unweighted annual release counts and displays the concentration of early releases from 2015 to 2017 alongside the distribution of qualifying releases across the entire sample period.



(a) Raw Absolute Innovation Value (winsorized at 99th pct)



(b) Log-Transformed Value with Normal Overlay

Figure 3. Distribution of Cohort-Level Intensity Weights

This figure plots the cohort-level distribution of the absolute Kogan (2017) innovation value used as the analytic weight in the stacked Difference-in-Differences (DiD) estimation. Panel (a) shows the raw distribution of the absolute intensity, winsorized at the 99th percentile for display clarity. Panel (b) displays the log-transformed distribution with a normal-density overlay.

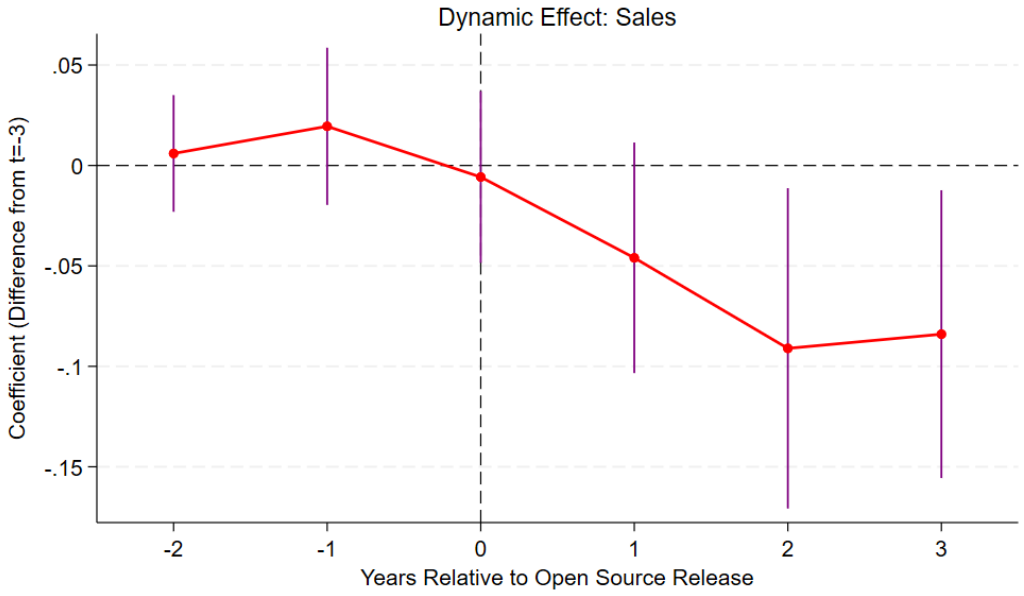
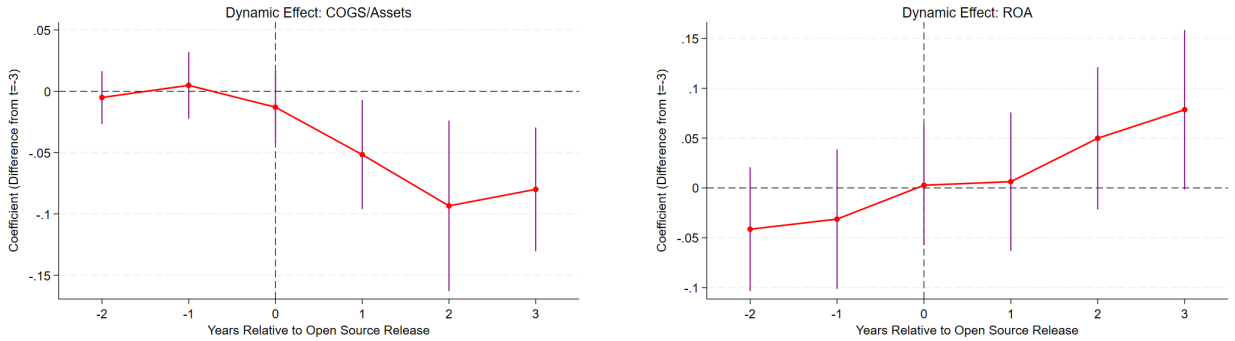


Figure 4. Dynamic Treatment Effect on Peer Firm Sales

This figure plots the dynamic treatment effect of focal firms' open-source releases on peer firms' Sales normalized by total assets. The estimation utilizes the full sample with the absolute cohort-level Kogan (2017) innovation value as an analytic weight. The specification includes Cohort-Firm and Cohort-Time fixed effects, firm-level controls (Market Capitalization, Market-to-Book Ratio, Leverage), and firm-level clustered standard errors. The x-axis indicates the years relative to the focal firm's open-source release (Year 0). The y-axis reports the estimated coefficients as the difference from the base year ($t = -3$), accompanied by their 95% confidence intervals.



(a) Cost of Goods Sold (COGS)

(b) Return on Assets (ROA)

Figure 5. Dynamic Treatment Effect on Peer Firm Cost and Profitability

This figure presents the dynamic treatment effects on the cost and profitability measures of peer firms. Panel (a) plots the estimated effect on Cost of Goods Sold (COGS) normalized by total assets, and Panel (b) plots the effect on Return on Assets (ROA). For both panels, the x-axis indicates the years relative to the focal firm's open-source release (Year 0), and the y-axis shows the estimated coefficients relative to the base year ($t = -3$), accompanied by their 95% confidence intervals. All estimates are derived from the weighted DiD specification incorporating Cohort-Firm and Cohort-Time fixed effects, firm-level controls, and firm-level clustered standard errors.

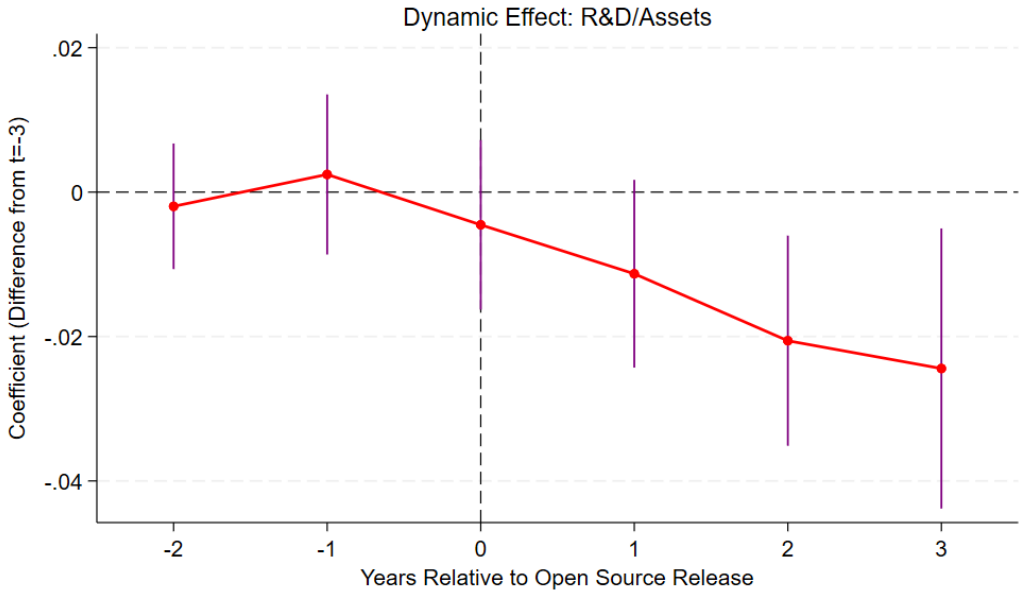
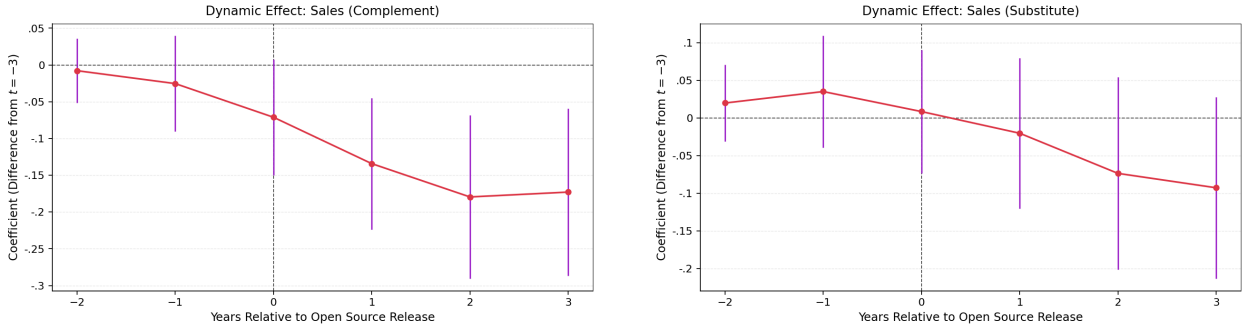


Figure 6. Dynamic Treatment Effect on Peer Firm R&D

This figure plots the dynamic treatment effect of focal firms' open-source releases on peer firms' Research and Development (R&D) expenditure normalized by total assets. The specification includes Cohort-Firm and Cohort-Time fixed effects, firm-level controls (Market Capitalization, Market-to-Book Ratio, Leverage), firm-level clustered standard errors, and the absolute cohort-level Kogan (2017) innovation value as the analytic weight. The x-axis indicates the years relative to the focal firm's open-source release (Year 0). The y-axis reports the estimated coefficients as the difference from the base year ($t = -3$), accompanied by their 95% confidence intervals.



(a) Complement Peers

(b) Substitute Peers

Figure 7. Heterogeneous Dynamic Effect on Sales: Complement vs. Substitute Peers
This figure plots the heterogeneous dynamic treatment effects on Sales for peer firms, partitioned by CPC technology overlap (TF-IDF cosine similarity between the peer's and the releasing firm's patent-CPC profiles) and industry-adjusted stock-return correlation with the focal firm over the pre-event window $[t-250, t-30]$. Panel (a) displays the effect on Complement peers, defined by below-median CPC overlap (an adjacent technology layer) and positive return co-movement. Panel (b) displays the effect on Substitute peers, defined by above-median CPC overlap (a direct technological rival) and negative return co-movement. The dynamic Difference-in-Differences estimates are weighted by the absolute cohort-level Kogan (2017) innovation value. The x-axis indicates years relative to the release (Year 0), and the y-axis reports the coefficients relative to the base year ($t = -3$), accompanied by 95% confidence intervals.

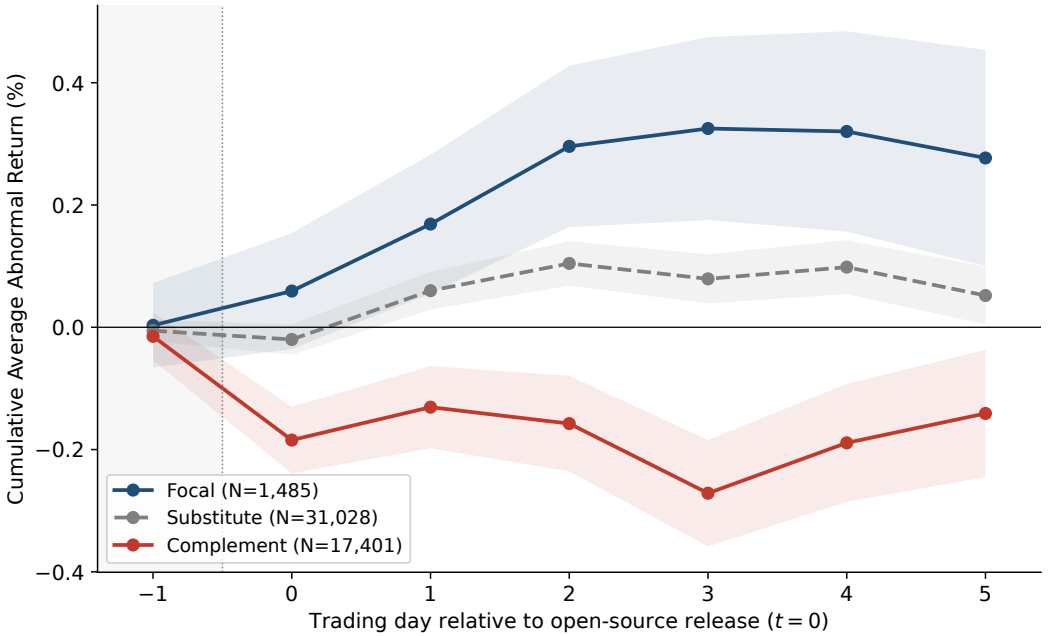


Figure 8. Market Reaction to Open-Source Releases by Peer Type

This figure plots the value-weighted cumulative average abnormal return for focal firms, substitute peers, and complement peers around open-source release announcements. The sample is the full set of approximately 1,700 open-source releases. The horizontal axis shows trading days relative to the release day ($t = 0$). The placebo day $t = -1$ is statistically indistinguishable from zero for all three groups. Abnormal returns are estimated from the Fama-French four-factor (FF4) model with factor loadings fit on a 150-day estimation window ending 30 days before each event. Firms are weighted by pre-event market capitalization. Shaded bands represent 95% confidence intervals constructed from the effective-sample-size weighted cross-sectional variance. Substitutes and complements are partitioned by CPC technology overlap, the asset-price classification described in Section 6.3, which excludes the return-co-movement criterion to keep the partition independent of the return outcome.

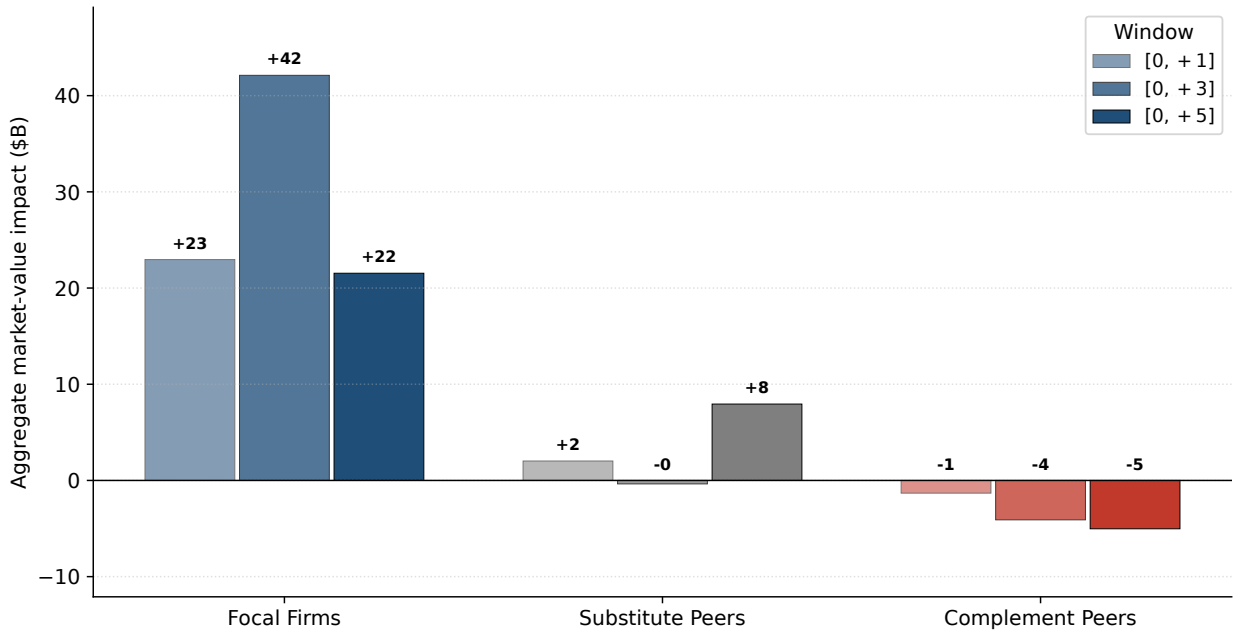


Figure 9. Aggregate Market-Value Impact by Peer Type

This figure reports the aggregate dollar-value impact of open-source software releases on focal firms, substitute peers, and complement peers for the full sample of approximately 1,700 open-source releases. Each bar shows the cross-sectional sum of the firm-level Fama-French four-factor CAR over the indicated post-event window multiplied by pre-event market capitalization. Three post-event windows are shown: $[0, +1]$, $[0, +3]$, and $[0, +5]$. Complement and substitute peers are defined by CPC technological overlap and the sign of pre-event return co-movement with the focal firm, as described in Section 6.1.

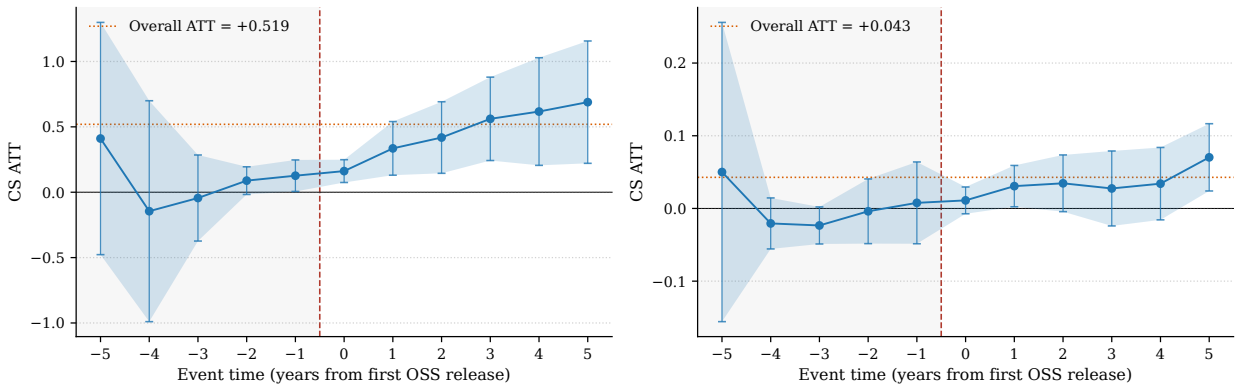


Figure 10. Focal Firm Performance Around Open-Source Release

This figure plots the Callaway–Sant'Anna doubly-robust ATT for the two headline financial outcomes as a function of event time relative to each focal firm's first eligible open-source release. The left panel reports Log Sales and the right panel reports Return on Assets. The estimator uses the never-treated propensity-score-matched controls as the comparison group and aggregates clean 2×2 cohort-time comparisons. Vertical bars are 95% pointwise confidence intervals; the shaded band on the left of each panel covers the pre-release period; the orange dashed horizontal line reports the overall ATT across the post-release window. The sample is the 15-focal PSM panel described in Table 9.

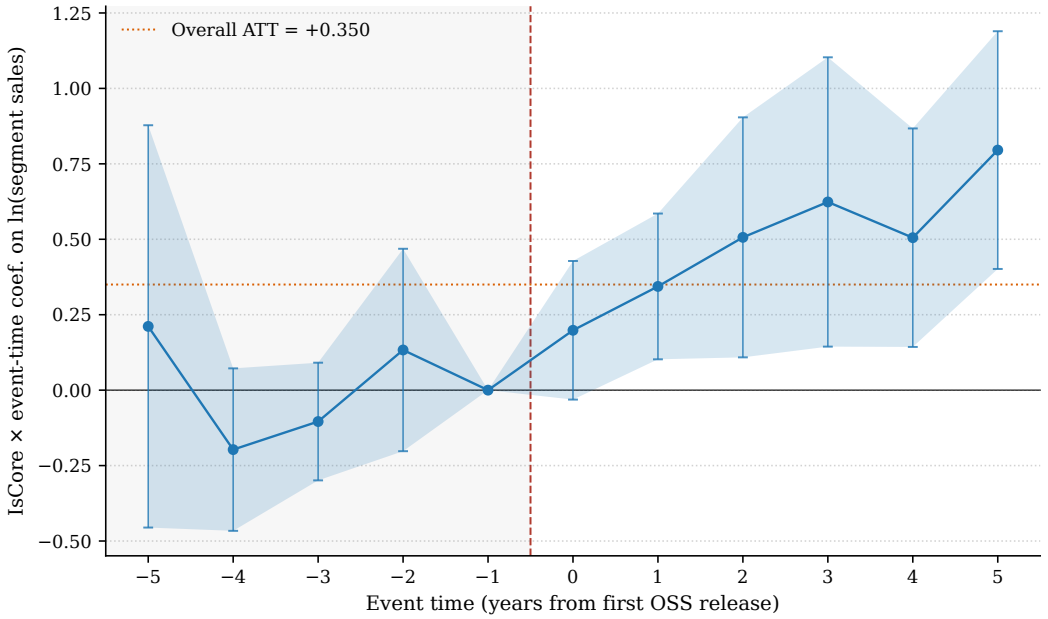


Figure 11. Focal Firm Core-Segment Demand Around Open-Source Release

This figure plots the within-firm-year event study of focal-firm Core-segment demand around the firm's first eligible open-source release. Each point is the coefficient on the interaction of the segment's Core indicator with an event-time indicator in a regression of ln(segment sales) on firm×year and firm×segment fixed effects, estimated on the 8-focal segment panel described in Table 10. Event time is measured in fiscal years relative to the release year, and the base year is $t = -1$. Endpoints are binned at ± 5. Vertical bars are 95% confidence intervals from standard errors clustered at the firm×segment level; the shaded band covers the pre-release period; the orange dotted horizontal line reports the pooled post-release effect.

Table 1. Descriptive Summary of the Full Open-Source Release Sample

This table presents descriptive statistics for the 1,671 open-source releases utilized in the empirical analysis. Panel A reports summary statistics by quartiles of the Kogan (2017) innovation value, defined as the focal firm's market-value reaction (in millions of U.S. dollars) attributable to the release. The columns report the number of releases (*N*), the average cumulative abnormal return around the release window (CAR, %), the average intensity ($M), average GitHub Stars one year post-release, and the releasing firm's average market capitalization ($B) on the release date. Panel B reports the top 15 of the 20 focal firms ranked by the number of releases, alongside their total intensity ($M), average CAR (%), and total GitHub Stars accumulated across their respective releases; the 15 firms shown account for 1,663 of the 1,671 releases. The sample period spans from 2015 to 2024.

Panel A: Summary by Kogan Intensity Quartiles

Quartile	N	Avg. CAR (%)	Avg. Intensity ($M)	Avg. Stars	Avg. Mkt Cap ($B)
Q1 (Lowest)	418	-2.09	-16,974.9	5,386	980.7
Q2	418	-0.48	-1,120.1	5,428	338.9
Q3	417	+1.27	+2,833.4	8,087	406.0
Q4 (Highest)	418	+2.53	+19,906.1	5,208	1,010.8

Panel B: Top 15 Focal Firms by Release Volume

Company Name	# of Repos	Total Intensity ($M)	Avg. CAR (%)	Total Stars
Microsoft	534	+1,105,058	+0.30	2,966,059
Alphabet	469	+31,752	+0.09	3,435,776
Meta	291	+139,906	+0.31	2,220,525
Amazon	101	-281,229	-0.24	309,247
Nvidia	100	+316,684	+0.52	378,405
Apple	52	+450,913	+0.58	303,754
Netflix	43	+51,672	+2.88	208,381
Salesforce	22	-2,335	-0.04	54,539
IBM	16	+1,446	+0.06	31,865
Oracle	10	+3,506	+0.15	44,552
Adobe	8	-603	+0.23	44,744
Cisco	6	+2,827	+0.34	20,536
Tesla	4	+116,796	+2.26	7,607
Datadog	4	+2,384	+1.85	17,905
Elastic	3	+284	+1.42	5,124

Table 2. Summary Statistics: Treated vs. Matched Control

This table presents pre-event year ($t = -1$) summary statistics for the main firm-year level variables used in the empirical analysis. Panel A reports the statistics for the treated cohort, while Panel B reports the matched-control peer firms. Financial ratios (Sales, COGS, R&D) are normalized by total assets. Market-to-Book Ratio is the market value of equity divided by the book value of equity. Market Capitalization and Total Assets are expressed in billions of U.S. dollars. Counts (N) refer to cohort-firm observations. The sample period spans 2015–2024.

	N	Mean	Std. Dev.	p10	Median	p90
Panel A: Treated Sample						
Sales	86,614	0.783	0.588	0.317	0.629	1.440
COGS	86,614	0.342	0.465	0.052	0.191	0.805
R&D	86,616	0.131	0.120	0.018	0.107	0.246
ROA	86,614	-0.087	0.629	-0.276	-0.006	0.119
Market Cap ($ Billions)	86,616	32.293	131.185	0.091	2.388	47.320
Market-to-Book	86,616	5.516	122.736	1.074	4.379	19.202
Leverage	86,616	0.192	0.194	0.000	0.149	0.450
Total Assets ($ Billions)	86,616	12.895	44.720	0.065	0.965	18.769
Panel B: Matched Control Sample						
Sales	66,668	0.664	0.430	0.246	0.584	1.132
COGS	66,668	0.307	0.340	0.050	0.184	0.712
R&D	66,668	0.080	0.087	0.000	0.058	0.188
ROA	66,668	-0.071	0.463	-0.293	0.010	0.126
Market Cap ($ Billions)	66,668	23.304	126.379	0.068	2.007	30.297
Market-to-Book	66,668	6.555	59.683	0.882	3.614	16.479
Leverage	66,668	0.201	0.251	0.000	0.140	0.458
Total Assets ($ Billions)	66,668	8.355	31.907	0.060	1.033	14.220

Table 3. The Impact of Big Tech Open-Source Release on Peer Firms' Sales

This table presents the results of a stacked Difference-in-Differences (DiD) estimation of the effect of focal firms' open-source releases on the sales of peer firms. Column (1) presents a baseline OLS specification without fixed effects or controls. Column (2) incorporates Cohort-Firm and Cohort-Time fixed effects along with firm-level controls. Column (3) further weights the observations using the absolute value of the cohort-level Kogan (2017) innovation value to emphasize economically consequential releases and reduce attenuation from marginal cohorts. Control variables include Market Capitalization, Market-to-Book (M/B) Ratio, and Leverage. Standard errors are clustered at the firm level in Columns (2) and (3). t-statistics are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample period spans from 2015 to 2024.

	(1)	(2)	(3)
Dependent Variable =	Sales	Sales	Sales
Treated × Post	-0.049***	-0.046**	-0.060***
	(-21.58)	(-2.39)	(-2.69)
Market Cap		0.091***	0.081**
		(2.69)	(2.28)
M/B Ratio		-0.002	-0.002
		(-0.88)	(-1.51)
Leverage		0.069*	0.071**
		(1.91)	(2.07)
Adjusted R^2	0.009	0.715	0.735
Cohort-Firm FE	No	Yes	Yes
Cohort-Time FE	No	Yes	Yes
Clustering	None	Firm	Firm
Weight	No	No	Yes
Mean of Dep. Var.	0.728	0.728	0.728
Observations	974,473	974,277	974,277

Table 4. The Impact of Big Tech Open-Source Release on Peer Firms' Profitability and Costs
This table presents the results of a stacked Difference-in-Differences (DiD) estimation of the effect of focal firms' open-source releases on the costs and profitability of peer firms. The dependent variables are Cost of Goods Sold (COGS) and Return on Assets (ROA), both normalized by total assets. Columns (1) and (3) present a baseline OLS specification without fixed effects or controls. Columns (2) and (4) incorporate Cohort-Firm and Cohort-Time fixed effects along with firm-level controls, and weight the observations using the absolute value of the cohort-level Kogan (2017) innovation value to emphasize economically consequential releases and reduce attenuation from marginal cohorts. Control variables include Market Capitalization, Market-to-Book (M/B) Ratio, and Leverage. Standard errors are clustered at the firm level in Columns (2) and (4). t-statistics are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample period spans from 2015 to 2024.

	Cost		Profitability	
	(1)	(2)	(3)	(4)
Dependent Variable =	COGS	COGS	ROA	ROA
Treated × Post	-0.045^{***}	-0.054^{***}	0.018^{***}	0.053^{**}
	(-25.58)	(-3.00)	(7.46)	(2.00)
Market Cap		0.017		-0.014
		(1.46)		(-0.36)
M/B Ratio		-0.002^{*}		-0.000
		(-1.72)		(-0.10)
Leverage		0.034		-1.066^{***}
		(0.98)		(-18.00)
Adjusted R^2	0.001	0.759	0.000	0.589
Cohort-Firm FE	No	Yes	No	Yes
Cohort-Time FE	No	Yes	No	Yes
Clustering	None	Firm	None	Firm
Weight	No	Yes	No	Yes
Mean of Dep. Var.	0.328	0.328	-0.064	-0.064
Observations	974,473	974,277	974,473	974,277

Table 5. The Impact of Big Tech Open-Source Releases on Peer Firms' GitHub Activity

This table presents stacked Difference-in-Differences (DiD) estimates of the effect of focal firms' open-source releases on the GitHub engagement of peer firms, using the full sample of approximately 1,700 release cohorts. Observations are weighted by the cohort-level absolute Kogan (2017) innovation value to anchor the estimates on the most economically influential releases. Columns (1) to (3) report results for per-developer activity (total, active, and passive contributions); column (4) reports the effect on the size of the active developer base. *Active* contributions include pushes, pull requests, issues, issue comments, and forks; *Passive* contributions include stars and watches. All specifications include firm-level controls (Market Capitalization, Market-to-Book Ratio, Leverage), Cohort-Firm and Cohort-Time fixed effects, and firm-clustered standard errors. The unweighted variant is reported in Internet Appendix Table IA.4, and the corresponding dynamic event-study is reported in Internet Appendix Figure IA1. t-statistics are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample period spans 2015–2024.

Dependent Variable =	Activity per Developer			Developer Base
	(1) Any Work per Dev.	(2) Active Contrib.	(3) Passive Contrib.	(4) Ln(Working Devs + 1)
Treated × Post	1.145***	1.134**	0.011	0.141***
	(2.60)	(2.57)	(0.41)	(5.32)
Adjusted R^2	0.743	0.740	0.422	0.839
Controls	Yes	Yes	Yes	Yes
Cohort-Firm FE	Yes	Yes	Yes	Yes
Cohort-Time FE	Yes	Yes	Yes	Yes
Clustering	Firm	Firm	Firm	Firm
Weight	Yes	Yes	Yes	Yes
Mean of Dep. Var.	3.122	2.801	0.321	0.480
Observations	974,385	974,385	974,385	974,385

Table 6. The Impact of Big Tech Open-Source Releases on Peer Firms' R&D Investment

This table presents the results of a stacked Difference-in-Differences (DiD) estimation of the effect of focal firms' open-source releases on the research and development (R&D) expenditures of peer firms, scaled by total assets. Column (1) presents a baseline OLS specification without fixed effects or controls. Column (2) incorporates Cohort-Firm and Cohort-Time fixed effects along with firm-level controls. Column (3) further weights the observations using the absolute value of the cohort-level Kogan (2017) innovation value to emphasize economically consequential releases and reduce attenuation from marginal cohorts. Control variables include Market Capitalization, Market-to-Book (M/B) Ratio, and Leverage. Standard errors are clustered at the firm level in Columns (2) and (3). t-statistics are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample period spans from 2015 to 2024. Companion estimates on peer firms' patent portfolio age (overall and exploratory) are reported in Internet Appendix Table IA.5.

	(1)	(2)	(3)
Dependent Variable =	R&D	R&D	R&D
Treated × Post	-0.006^{***}	-0.012^{***}	-0.014^{***}
	(-10.54)	(-2.75)	(-2.75)
Market Cap		0.018^{**}	0.013^{**}
		(2.26)	(2.09)
M/B Ratio		0.000	0.001^{**}
		(0.24)	(2.49)
Leverage		0.219^{***}	0.262^{***}
		(3.56)	(7.37)
Adjusted R^2	0.009	0.732	0.788
Cohort-Firm FE	No	Yes	Yes
Cohort-Time FE	No	Yes	Yes
Clustering	None	Firm	Firm
Weight	No	No	Yes
Mean of Dep. Var.	0.107	0.107	0.107
Observations	993,135	992,931	992,931

Table 7. Heterogeneous Effects on Substitute vs. Complement Peers

This table reports stacked Difference-in-Differences (DiD) estimates of the effect of focal firms' open-source releases on peer firms' Sales and Cost of Goods Sold (COGS), separately for Substitute and Complement peers under a two-dimensional classification. The partition intersects technology overlap with stock-return co-movement and retains only the peers on which both signals agree. A Substitute has above-median CPC cosine similarity with the releasing firm (high technology overlap, a direct technological rival) together with negative pre-event industry-adjusted return co-movement; a Complement has below-median CPC overlap (an adjacent technology layer) together with positive return co-movement. Peers for which the two signals disagree are dropped as ambiguous, which sharpens the economic contrast at the cost of sample size. The dependent variables are normalized by total assets. Each column compares the relevant treated-peer group against the common matched-control pool, so the matched controls are shared across the Substitute and Complement columns; the two observation counts therefore overlap in their control rows and need not sum to the pooled sample. All specifications include Cohort-Firm and Cohort-Time fixed effects and firm-level controls, and observations are weighted by the absolute value of the cohort-level Kogan (2017) innovation value. Standard errors are clustered at the firm level. t-statistics are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample period spans from 2015 to 2024.

Dependent Variable =	Sales		COGS	
Peer Firm Type =	(1) Substitute	(2) Complement	(3) Substitute	(4) Complement
Treated × Post	-0.0545	-0.1211***	-0.0466*	-0.0765***
	(-1.60)	(-3.69)	(-1.82)	(-2.86)
Adjusted R^2	0.630	0.681	0.627	0.701
Controls	Yes	Yes	Yes	Yes
Cohort-Firm FE	Yes	Yes	Yes	Yes
Cohort-Time FE	Yes	Yes	Yes	Yes
Clustering	Firm	Firm	Firm	Firm
Weight	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.681	0.700	0.315	0.338
Observations	422,642	448,366	422,642	448,366

Table 8. Market Reaction to Open-Source Releases by Peer Type

This table reports value-weighted cumulative abnormal returns around open-source release announcements for the full sample of approximately 1,700 open-source releases. Window $[-1,0)$ covers the single trading day immediately before the release ($t=-1$) and serves as a within-window placebo. Windows $[0,+k]$ for $k \in \{1,3,5\}$ report the cumulative CAR from the release day through $+k$ trading days. Abnormal returns are estimated from the Fama-French four-factor (FF4) model with factor loadings fit on a 150-day estimation window ending 30 days before each event. Firms are weighted by pre-event market capitalization. Substitutes and complements are partitioned by CPC technology overlap: a Substitute has above-median CPC cosine similarity with the focal firm, and a Complement has below-median overlap together with membership in the focal firm's broad SIC2 sector. The return-co-movement criterion used in the product-market partition of Section 6.1 is excluded from this asset-price test to keep the classification independent of the return outcome. t-statistics in parentheses are computed from the effective-sample-size weighted cross-sectional variance. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels.

	$[-1,0)$	$[0,+1]$	$[0,+3]$	$[0,+5]$
Focal	$+0.003\%$	$+0.166\%^{***}$	$+0.322\%^{***}$	$+0.274\%^{***}$
N=1,485	$(+0.09)$	$(+3.67)$	$(+5.01)$	$(+3.44)$
Substitute	-0.006%	$+0.065\%^{***}$	$+0.085\%^{***}$	$+0.058\%^{***}$
N=31,028	(-0.63)	$(+5.09)$	$(+4.59)$	$(+2.63)$
Complement	-0.015%	$-0.116\%^{***}$	$-0.257\%^{***}$	$-0.126\%^{**}$
N=17,401	(-0.74)	(-4.09)	(-6.13)	(-2.55)

Table 9. Focal Firm Performance After Open-Source Release

This table reports the focal-firm Difference-in-Differences estimates for the two headline financial outcomes, Log Sales and Return on Assets, under two estimators on a propensity-score-matched sample of 15 focal firms with their 5-nearest-neighbor controls. Columns (1) and (3) report the two-way fixed-effect (TWFE) estimate of equation (2) with firm and fiscal-year fixed effects and firm-clustered standard errors. Columns (2) and (4) report the Callaway and Sant'Anna (2021) doubly-robust ATT estimator with a never-treated control group, which weights only clean 2×2 cohort-time comparisons. The 15 focal firms (Adobe, Alphabet, Amazon, Apple, Cisco, IBM, Meta, Microsoft, Netflix, Nvidia, Oracle, Qualcomm, Salesforce, Tesla, and Twilio) are matched to 53 unique controls drawn from the Compustat tech-firm pool (SIC 73, 35, 36, 38) via 5-nearest-neighbor propensity-score matching on pre-release ln(total assets), ln(sales), ROA, market-to-book, and leverage, with exact two-digit SIC matching and a 0.25-standard-deviation caliper. Matching is with replacement, and the estimation panel stacks each focal cohort with its matched controls, so a control firm that serves as a comparison for more than one focal cohort enters once per cohort to which it is matched. The 53 unique controls therefore contribute more than 53 firm-year series, and the 1,163 stacked firm-cohort-year observations exceed the 872 unique firm-years in the pooled sample; the firm-clustered standard errors account for these repeated control appearances. The panel spans fiscal years 2010–2023. *t*-statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)
Dependent Variable =	Log Sales		ROA	
Focal × Post	$+0.838^{***}$	$+0.519^{***}$	$+0.042^{*}$	$+0.043^{***}$
	(3.30)	(3.18)	(1.96)	(2.85)
Adjusted R^2	0.306	—	0.055	—
Firm FE	Yes	—	Yes	—
Year FE	Yes	—	Yes	—
Estimator	TWFE	CS-DR	TWFE	CS-DR
Observations	1,163	1,163	1,163	1,163

Table 10. Focal Firm Core-Segment Demand After Open-Source Release

This table reports a within-firm-year difference-in-differences estimate of how a focal firm's open-source releases reallocate demand toward its core business segment. The outcome is ln(segment sales) from Compustat Segment data. The panel covers the 8 focal firms (Adobe, Alphabet, IBM, Meta, Microsoft, Nvidia, Oracle, and Qualcomm) from the focal-firm panel in Table 9 that report both Core and Non-Core business segments, so that the within-firm Core-versus-Non-Core contrast is well defined; the remaining focal firms either lack a multi-segment breakdown or report only a single non-core segment, as detailed in Appendix Section B.3. Each business segment is hand-classified as Core or Non-Core from the focal firm's 10-K disclosure, with the full mapping reported in Internet Appendix Table IA.6. *IsCore* × *Post* equals one for Core segments in the fiscal years on or after the focal firm's first eligible release. Columns (1) and (2) use the full sample (fiscal years 2010–2023); columns (3) and (4) restrict to fiscal years 2014–2023. All specifications include firm×year and firm×segment fixed effects, so identification comes purely from within-firm-year differential growth of Core relative to Non-Core segments around the release, as specified in equation (3). Columns (1) and (3) report unadjusted standard errors; columns (2) and (4) cluster standard errors at the firm×segment level. The event-study counterpart is plotted in Figure 11. t-statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)
Dependent Variable =	ln(segment sales)			
Sample period =	Full sample (2010–2023)		Post-2014 (2014–2023)	
IsCore × *Post*	$+0.350^{***}$	$+0.350^{***}$	$+0.471^{***}$	$+0.471^{***}$
	(+3.79)	(+3.01)	(+4.27)	(+3.38)
Adjusted R^2	0.496	0.496	0.339	0.339
Firm × Year FE	Yes	Yes	Yes	Yes
Firm × Segment FE	Yes	Yes	Yes	Yes
Clustering	None	Firm × Seg	None	Firm × Seg
Observations	739	739	493	493

Internet Appendix for:
"Commoditize Your Complement: Strategic Open-Source Releases and the Cannibalization of Rival Firms"

A Theoretical Framework: A Simple Multi-Segment Model

This section sets out a deliberately simple model in which strategic open-source release is the commoditization of one segment of a multi-segment firm's product portfolio. The model is a reduced-form device meant to fix ideas and to organize the empirical analysis around the distinction between substitute and complement peers that the paper turns on, not to provide a structural account of the open-source decision; the paper's central contribution is empirical. The construction proceeds from four primitives: the focal firm's three-segment profit structure, the open-source license as a commitment device, the peer firm's exposure vector across the same three segments, and the four channels of the focal firm's release decision. From these primitives the framework generates five predictions. The first concerns the focal firm's own valuation. The second, which is the paper's centerpiece, is an asymmetric sign prediction across peer types that no alternative rationalization of the release reproduces. The remaining three characterize the mechanism chain through which peer outcomes adjust on real-economy margins.

A.1 A Multi-Segment Firm as the Primitive Unit of Analysis

The analytical object is not the firm but the multi-segment firm, operating across three logically distinct product markets indexed by $j \in \{A, B, C\}$. Each segment has its own demand curve, its own set of rivals, and its own strategic role in the portfolio. Segment A is the profit center, generating the monopoly or quasi-monopoly rents that the firm has every incentive to protect. Advertising inventory for Meta, search advertising for Alphabet, and enterprise cloud for Microsoft are canonical examples. Segment B is a separate market the focal firm also occupies, where entry by a rival threatens the A-segment boundary; because it is additively separable and a release in C leaves it unaffected, B plays no role in the release decision or in the peer cross-section developed below. Segment C is a complement market to A, in which higher consumption raises the marginal value of A through a cross-market demand linkage. In segment C the optimal posture is commoditization rather than rent extraction, because cheaper C-goods expand A-market demand. This three-segment partition restates the distinction that Adner and Kapoor (2010) draw between upstream components and downstream complements in innovation ecosystems, applied here to the firm's own internal portfolio rather than to an ecosystem of separate producers.

Let focal firm f earn profit

$$\Pi_f = \pi_{f,A}(p_A, Q_C) + \pi_{f,B}(p_B) + \pi_{f,C}(p_C, p_{-C}), \tag{A-4}$$

where p_j is the price in segment j, Q_j is aggregate quantity, and p_{-C} denotes rival prices in segment C. To give the cross-segment linkage explicit content, the model adopts the cross-market externality structure of Parker and Van Alstyne (2005), specialized from their two-sided platform of distinct customer groups to a within-firm complement-to-profit-center linkage. Segment A faces the inverse demand

$$p_A = a_A - b_A Q_A + \theta\, Q_C, \qquad a_A, b_A > 0,\ \ \theta \geq 0, \tag{A-5}$$

in which θQ_C is the additional willingness to pay that complement consumption confers on the profit center, the port of the cross-side network effect in Parker and Van Alstyne (2005). For the focal firm's own problem segment C is a standalone market with inverse demand $p_C = a_C - b_C Q_C$ and marginal cost normalized to zero. Profit-center rent is then $\pi_{f,A} = (a_A - b_A Q_A + \theta Q_C - c_A)\, Q_A$, so that

$$\frac{\partial \pi_{f,A}}{\partial Q_C} = \theta\, Q_A \ \geq\ 0, \tag{A1}$$

strictly positive whenever $\theta > 0$. The cross-subsidy premise of the commoditize-your-complement heuristic, stated informally by Spolsky (2002), is therefore not an independent assumption but a consequence of a positive cross-market externality. This is the classic economics of complements: Cournot (1838) showed that the owner of one of two complementary goods gains when the price of the other falls, so a firm that controls the profit center has a strict incentive to drive the complementary layer's price toward marginal cost, the systems logic synthesized by Varian (2004). The restriction $\theta > 0$ is labeled (A1). Because the cross-partial scales with Q_A, the linkage is mechanically stronger for firms with a larger profit center, the formal counterpart of the profit-center-dominance condition in Proposition 1.

Peer firm r is characterized by an exposure vector $(s_{r,A}, s_{r,B}, s_{r,C}) \in [0,1]^3$ with $\sum_j s_{r,j} = 1$, representing the fraction of the peer's revenue derived from each segment. A peer is not defined by an external label but by where its revenue lives. This representation makes the substitute-versus-complement distinction that organizes the empirical results an observable projection of the peer's segment exposure rather than an analyst-imposed taxonomy, a property we return to in Section A.3.

A.2 Strategic Release as Commoditization of Segment C

The focal firm faces a discrete choice. It either extracts rent in segment C through proprietary pricing or releases a high-quality asset Ω in segment C under an open-source license at consumer price zero in perpetuity, financed out of segment A's rents. Denote this choice $OSS_f \in \{0,1\}$.

Three features of the open-source license separate this choice from an ordinary price cut. First, the release is irreversible: once code is distributed under an open-source license, no actor including the releasing firm can return it to proprietary status, because any attempted closure is defeated by forks of the released version. Second, the subsidy flows across segments rather than within. The cost of providing Ω is absorbed by segment A profits, not by an expectation that segment C prices will eventually rise to recoup. Third, the release is a technical standard-setting act in addition to a pricing act. Adoption of Ω as common infrastructure creates network effects and switching costs that persist beyond the release announcement. The combination of these features makes the standard predatory-pricing hold-up logic inapplicable. A competitor cannot wait out a rival whose subsidy never exhausts because the subsidy does not come from the same pool the competitor is defending. This commitment structure, inherited from the legal form of open-source licensing rather than from discretionary managerial behavior, is what separates the strategy from an ordinary price war.

The change in focal profit from releasing Ω decomposes into four channels:

$$\Delta\Pi_f \;=\; \underbrace{\frac{\partial \pi_{f,A}}{\partial Q_C}\frac{\partial Q_C}{\partial OSS_f}}_{\text{(i) cross-market subsidy}} \;-\; \underbrace{\left|\frac{\partial \pi_{f,C}}{\partial OSS_f}\right|}_{\text{own } C\text{-rent loss}} \;+\; \underbrace{\mathbb{E}\left[\sum_r \left(-\Delta\pi_{r,C}\right)\phi_{fr}\right]}_{\text{(ii) peer redirection}} \;+\; \underbrace{\lambda_3}_{\text{(iii) standards}} \;+\; \underbrace{\lambda_4}_{\text{(iv) talent}}. \tag{A-6}$$

Channel (i) is the cross-market subsidy gain that operates through assumption (A1). Channel (ii) is the focal firm's captured benefit from rival redirection: a release that compresses an exposed rival's C-segment rent, $\Delta\pi_{r,C} \le 0$, diverts to the focal firm a share $\phi_{fr} \ge 0$ of that forgone rent through downstream complementarities, so the channel equals $\mathbb{E}\left[\sum_r(-\Delta\pi_{r,C})\,\phi_{fr}\right] \ge 0$ and is distinct from the rival's own (negative) profit change. Channel (iii), worth $\lambda_3 \ge 0$, is the option value of making Ω the technical standard, which in the tradition of Farrell and Saloner (1985) and Katz and Shapiro (1985) locks future entrants onto the focal firm's interface. Channel (iv), worth $\lambda_4 \ge 0$, is the option value of drawing outside developer contribution onto the focal firm's technology trajectory. Release is optimal when the sum of the four positive channels dominates the direct C-rent loss.

Channels (ii), (iii), and (iv) are non-negative and are collected into a single option term $\Lambda \equiv \lambda_3 + \lambda_4 + \mathbb{E}\left[\sum_r \Delta\pi_{r,C}\,\phi_{fr}\right] \ge 0$, kept in reduced form because the asymmetric peer prediction developed below does not depend on their internal structure. Channel (i), by contrast, is pinned down by the demand primitives in (A-5). Optimizing profit-center output, $\partial\pi_{f,A}/\partial Q_A = 0$, gives $Q_A^*(Q_C) = (a_A - c_A + \theta Q_C)/(2b_A)$ and the optimized rent

$$\pi_{f,A}^*(Q_C) = \frac{(m+\theta Q_C)^2}{4b_A}, \qquad m \equiv a_A - c_A, \tag{A-7}$$

which is increasing and convex in complement consumption. Proprietary operation of segment C

yields the monopoly quantity $Q_C^0 = a_C/(2b_C)$ and own rent $\pi_{f,C}^0 = a_C^2/(4b_C)$; the irreversible release sets $p_C = 0$ and expands consumption to $Q_C^1 = a_C/b_C$, so that $\Delta Q_C \equiv Q_C^1 - Q_C^0 = a_C/(2b_C) > 0$. The cross-market subsidy of channel (i) is then the exact quantity

$$\Delta\pi_{f,A} = \pi_{f,A}^*(Q_C^1) - \pi_{f,A}^*(Q_C^0) = \frac{\theta\,\Delta Q_C\left[2m + \theta(Q_C^1 + Q_C^0)\right]}{4b_A}. \tag{A-8}$$

Proposition 1 (Release under *A*-dominance): *If $\pi_{f,A}$ is large relative to $\pi_{f,C}$ and $\partial\pi_{f,A}/\partial Q_C$ is sufficiently positive, then $OSS_f = 1$ is optimal even when $\partial\pi_{f,C}/\partial OSS_f < 0$.*

Derivation. Releasing is optimal when $\Delta\Pi_f = \Delta\pi_{f,A} - \pi_{f,C}^0 + \Lambda \geq 0$. Substituting (A-8) and the regime quantities, $OSS_f = 1$ is optimal if and only if

$$\underbrace{\frac{\theta\,\Delta Q_C\left[2m + \theta(Q_C^1 + Q_C^0)\right]}{4b_A}}_{\text{cross-market subsidy}} + \Lambda \;\geq\; \underbrace{\frac{a_C^2}{4b_C}}_{\text{forgone } C\text{-rent}}. \tag{A-9}$$

The left side increases in the cross-market externality θ and in the profit-center scale m, while the right side is the rent the firm forgoes by abandoning proprietary pricing in C. Holding the forgone rent fixed, the condition is met precisely when the profit center is large relative to the complement rent, which is the content of A-dominance. The signed comparative statics are on the cross-linkage and the profit-center scale: $\partial(\text{LHS})/\partial\theta > 0$ and $\partial(\text{LHS})/\partial m > 0$, so the firms that release are those with strong cross-linkage and large protected profit centers. The complement-demand primitives a_C and b_C do not deliver a signed prediction. They enter both sides of the condition in the same direction: the forgone monopoly rent is $\text{RHS} = a_C^2/(4b_C)$, while the cross-market subsidy scales with the induced expansion $\Delta Q_C = a_C/(2b_C)$, so $\partial(\text{LHS})/\partial a_C > 0$ and $\partial(\text{RHS})/\partial a_C > 0$ (both negative in b_C), leaving the net effect on the release condition ambiguous. □

The linear-quadratic demand is a transparency device, not a driver of the result. For a more general increasing $\pi_{f,A}(Q_C)$, the release condition retains the integral form $\int_{Q_C^0}^{Q_C^1}(\partial\pi_{f,A}/\partial Q_C)\,dQ_C + \Lambda \geq \pi_{f,C}^0$, so the qualitative release tradeoff between the cross-market gain and the forgone C-rent is preserved; the specific comparative-static signs reported below use the linear-quadratic structure.

Proposition 1 explains why the set of firms observed to release strategic open-source software is concentrated among technology incumbents whose profit base is dominated by a protected A-segment. Microsoft, Alphabet, Meta, Apple, and Nvidia have the subsidy pool to finance a permanent giveaway in an adjacent complement market. A firm whose profit base is itself concentrated in segment C has no equivalent pool from which to finance the release, and so does not release.

A corollary concerns the focal firm's announcement-window valuation. The firm undertakes

the release only when its net present value is non-negative under the firm's own information, so the positive contribution from channels (i), (iii), and (iv) outweighs the negative own-rent contribution and the release is, in expectation, value-increasing news. The announcement-window return is therefore non-negative in expectation, provided the release is not already anticipated by the market and conveys no offsetting information about costs, timing, or the firm's opportunity set; with these caveats the model places no hard lower bound on any single release's realized return.

Hypothesis 1 (Focal-Firm Valuation Gain): *Announcement of a strategic open-source release by a focal firm generates a non-negative cumulative abnormal return, on average, on the focal firm's equity in a narrow window around the release date.*

A.3 Peer Classification and the Asymmetric Response Prediction

The framework's sharpest prediction concerns the cross-section of peer firms. Unlike the focal firm's valuation, which the framework predicts to be non-negative on average, peer valuations can take either sign, and the distribution of outcomes is pinned down by how each peer's revenue is composed over the three segments.

Consider peer r with exposure vector $(s_{r,A}, s_{r,B}, s_{r,C})$. Define two statistics that summarize the peer's economic relation to the focal firm:

$$\sigma_r = s_{r,A}, \tag{A-10}$$

$$\gamma_r = s_{r,C}. \tag{A-11}$$

Segment-B exposure $s_{r,B}$ enters neither statistic. Although B is the market in which substitute entry would threaten the focal firm's moat, a release in C leaves B unaffected, so $s_{r,B}$ carries no predicted release effect and does not define a third peer category; the cross-section therefore turns on the single A-versus-C contrast, with $s_{r,B}$ absorbed in the residual $1 - \sigma_r - \gamma_r$. A substitute peer is accordingly a firm competing in the focal firm's profit center A, not a firm operating in B. A peer whose revenue is concentrated in segment A competes with the focal firm in the focal firm's own profit center. A peer whose revenue is concentrated in segment C supplies products that complement the focal firm's A-segment, operating in the same segment that Ω commoditizes. Adopting the terminology used throughout the empirical section, a peer with high σ_r and low γ_r is a *substitute peer* to the focal firm, because its products substitute for the focal firm's core offering. A peer with high γ_r and low σ_r is a *complement peer* to the focal firm, because its products sit alongside the focal firm's core offering in the consumption bundle. Oracle and Salesforce, whose enterprise software competes with Microsoft's Dynamics and Azure services, are substitute-peer examples. MongoDB and Confluent, whose data-layer products run on top of the hyperscalers'

core infrastructure, are complement-peer examples.

For a complement peer, the release directly compresses revenue. The peer sells in segment C, the same segment Ω now occupies at price zero. Consumers facing an Ω-quality free alternative and the peer's proprietary alternative at price $p_{r,C}$ with switching cost τ adopt Ω whenever $V_\Omega \geq V_{r,C} - p_{r,C} - \tau$ (Zhu and Zhou, 2012). Rising V_Ω through successive focal-firm updates compresses $p_{r,C}$ toward marginal cost, and the complement peer's C-segment revenue base contracts. Because $\sigma_r \approx 0$, the complement peer has no offsetting revenue in focal's A-adjacent markets. Its market valuation falls as the discounted present value of the revenue contraction.

For a substitute peer, the release produces no direct C-segment compression because $\gamma_r \approx 0$. The Q_C expansion induced by Ω raises the marginal value of segment A through the cross-partial in assumption (A1), which operates for the substitute peer's own A-segment revenue exactly as it does for the focal firm's. This effect is positive in sign, though typically modest in magnitude, because the substitute peer also competes with the focal firm in A and the focal firm's own expansion absorbs part of the Q_C-induced demand. The substitute peer's market valuation therefore adjusts by an amount that is weakly positive or statistically indistinguishable from zero.

Formally, write the peer's equity value as the exposure-weighted sum of segment values, so that a release perturbs only the A and C components,

$$\Delta V_r = \gamma_r \Delta v_C + \sigma_r \Delta v_A. \tag{A-12}$$

For a complement peer, the free asset Ω of quality V_Ω competes in segment C. With switching cost τ, the peer's sustainable price is $p_{r,C} = \max\{c, V_{r,C} - V_\Omega + \tau\}$, weakly decreasing in V_Ω, so the discounted change in C-segment value,

$$\Delta v_C = -\,\mathrm{PV}\big[\,\pi_{r,C}(p^0_{r,C}) - \pi_{r,C}(p_{r,C})\,\big] < 0, \tag{A-13}$$

is strictly negative once the release binds, $V_\Omega > V_{r,C} - p^0_{r,C} + \tau$. Because a complement peer has $\sigma_r \approx 0$,

$$\frac{\partial \Delta V_r}{\partial \gamma_r} = \Delta v_C < 0, \tag{A-14}$$

and the loss deepens in V_Ω, and, to the extent the Kogan intensity $|\kappa_r|$ proxies release quality, in that intensity as well, and shrinks in the switching cost τ.

For a substitute peer, $\gamma_r \approx 0$ removes the direct compression and the release reaches the peer only through its profit-center revenue, where two opposing forces act. The complement expansion raises the A-market intercept, of which the substitute peer captures a spillover share $\phi_s \in [0,1]$, a gain of order $\phi_s\,\theta\,\Delta Q_C$. Against it, the focal firm's optimal A-output expands with Q_C by (A-7), absorbing A-market demand from the substitute peer, a loss of order $\kappa\,\theta\,\Delta Q_C$ with $\kappa > 0$ the

competitive-displacement coefficient. The substitute peer's value change is therefore

$$\Delta V_{\text{sub}} = \sigma_r \, \theta \, \Delta Q_C \, (\phi_s - \kappa), \tag{A-15}$$

with sign equal to that of $\phi_s - \kappa$ and magnitude bounded by $\sigma_r \, \theta \, \Delta Q_C \max\{\phi_s, \kappa\}$. Two implications follow. First, the spillover and the absorption are both second-order cross-market adjustments, whereas the complement-peer compression in (A-13) is a first-order loss of own-segment rent, so $|\Delta V_{\text{sub}}| \ll |\Delta V_{\text{comp}}|$, and the net value-weighted substitute reaction is small and second-order — weakly positive or null — whatever the sign of $\phi_s - \kappa$. Second, the offset is a property of the value-weighted sum and need not hold outcome by outcome: on margins tied to head-to-head *A*-competition the absorption term dominates, predicting a weak negative response, while on margins tied to the commoditized *C*-segment the substitute peer is unexposed, predicting a null.

Proposition 2 (Asymmetric Response): *Announcement of $OSS_f = 1$ generates a strongly negative cumulative abnormal return on complement peers and a weakly positive or null cumulative abnormal return on substitute peers. The magnitude of the complement-peer loss is increasing in γ_r; the magnitude of the substitute-peer gain is increasing in σ_r and is bounded above by focal-firm absorption in segment A.*

No alternative rationalization of the release produces this asymmetry. A signaling account in which Ω is read as a quality signal about the focal firm predicts a uniform sign across peers, because all observers update the same prior. A philanthropy account predicts a concentrated valuation shock at the focal firm with no systematic peer effect. A Bertrand-predation story predicts uniform negative rival response across the full peer set, contradicting the observed null on substitute peers. A network-bootstrap account predicts mixed signs driven by network position rather than by segment exposure, and empirically would not align with a text-based product-market classifier. The asymmetric response is therefore a discriminating prediction, stated as a primary hypothesis rather than as an ex-post heterogeneity result:

Hypothesis 2 (Asymmetric Peer Valuation): *The cumulative abnormal return triggered by a strategic open-source release is concentrated on complement peers as a strong negative response and is weakly positive or null on substitute peers, second-order relative to the complement-peer loss. The asymmetry is a structural implication of the multi-segment firm's portfolio composition and does not require any assumption about the focal firm's targeting intent.*

Hypothesis 2 is the load-bearing prediction of the framework. It is tested directly in Section 6.3 using the focal firm's peer set partitioned into substitute and complement groups. The empirical operationalization, detailed in Section 6.1, projects the latent exposure pair (σ_r, γ_r) onto two observable signals: the sign of the peer's pre-event return co-movement with the focal firm and the

technological overlap between the peer's and the focal firm's patent portfolios. Complement peers are identified as peers whose pre-event returns co-move positively with the focal firm's and whose patent-CPC overlap with the focal firm is high. Substitute peers are identified as peers whose pre-event returns co-move negatively with the focal firm's. This projection loses information relative to a direct measure of segment revenue composition but has the advantage of using only pre-event data, preserving identification. Section A.4 derives the projection from the model's primitives, so that the partition is an implication of the framework rather than an empirical convention.

A.4 Identifying the Peer Types from Observable Signals

The partition in Section A.3 is stated in terms of the latent exposures (σ_r, γ_r), which are not directly observed. The empirical design recovers them by projecting each peer onto two observable signals: the technological overlap between the peer's and the focal firm's patent portfolios, denoted χ_r, and the peer's pre-event equity-return co-movement with the focal firm, denoted ρ_r. This subsection derives the mapping from (χ_r, ρ_r) to (σ_r, γ_r) from two measurement assumptions, so that the substitute-versus-complement partition is an implication of the model's primitives rather than an analyst-imposed rule. Both assumptions are deliberately minimal and require no re-estimation of the model.

(P1) Profit-center-weighted patenting. *The focal firm concentrates its patent portfolio in the technologies of its profit center A, which it defends, and patents little in the complement segment C, which it commoditizes. The focal firm's patent-class profile therefore points predominantly in the A-technology direction.*

(P2) Segment factor structure of returns. *Pre-event equity returns load on segment-level demand shocks ε_A and ε_C, with a peer's loading on each shock equal to its corresponding segment exposure. Through the cross-market linkage (A1), a positive complement-demand shock ε_C raises the rent of every firm with C-segment revenue and, via θ, the profit-center rent of every firm with A-segment revenue.*

Lemma 1 (Technology overlap identifies σ): *Under (P1), the patent-CPC overlap χ_r between peer r and the focal firm is increasing in σ_r. An above-median χ_r therefore selects high-σ (substitute) peers.*

Sketch. By (P1) the focal firm's patent-class profile is aligned with the *A*-technology direction. The overlap χ_r is the cosine similarity between peer *r*'s patent-class profile and the focal firm's, which equals the projection of peer *r*'s profile onto the *A*-direction. A peer that earns a larger share of its revenue in segment *A* patents more heavily in *A*-technologies, so the *A*-projection of its profile, and hence χ_r, rises in σ_r. The overlap is thus monotone in profit-center exposure and is informative about σ_r alone, independent of γ_r. □

Lemma 2 (Return co-movement identifies γ): *Under (P2) and (A1), the pre-event return co-movement ρ_r between peer r and the focal firm is increasing in γ_r. Positive co-movement selects high-γ (complement) peers; substitute peers, exposed to the focal firm mainly as profit-center rivals, co-move weakly or negatively.*

Sketch. By (P2) the focal firm's return loads positively on the complement-demand shock ε_C, directly and through the cross-market channel θ in (A1). A complement peer (γ_r high, $\sigma_r \approx 0$) shares this ε_C exposure, so its return and the focal firm's move together and $\rho_r > 0$. A substitute peer (σ_r high, $\gamma_r \approx 0$) shares the profit-center shock ε_A with the focal firm, but the two compete in segment A, where business stealing loads the peer's return on ε_A with a relative component opposite in sign to the focal firm's, the same competitive-displacement force that enters as κ in (A-15). The shared exposure is partly offset and ρ_r is pulled toward or below zero. Co-movement is therefore increasing in γ_r and decreasing in the A-competition that marks high-σ peers, so its sign tracks the complement-versus-substitute contrast. The data display this asymmetry: complement peers have mean pre-event co-movement $\bar{\rho} = 0.17$, while substitute peers have $\bar{\rho} = -0.07$, close to zero, consistent with the cross-market channel acting as the identifying signal on the complement side and the profit-center competition muting it on the substitute side. □

Corollary (the two-dimensional partition recovers the type ordering): *Combining Lemmas 1 and 2, the cell with χ_r above median and $\rho_r < 0$ collects high-σ peers and the cell with χ_r below median and $\rho_r > 0$ collects high-γ peers. Peers on the off-diagonal, whose two signals disagree, are noisy readings of the latent type and are dropped as ambiguous, which lowers classification error in (σ_r, γ_r) at the cost of sample size.*

The two-dimensional classifier used for the real-outcome tests is in this sense the projection of the model's latent types onto two derived signals rather than an ad-hoc sort. The same result disciplines the asset-price test. Return co-movement ρ_r is itself a pre-event return, so sorting peers on ρ_r and then measuring their announcement-window abnormal returns would condition the outcome on a function of past returns. The cumulative-abnormal-return test in Section 6.3 accordingly partitions peers on the technology signal χ_r alone, which by Lemma 1 identifies σ_r and is constructed entirely from patent data independent of returns. The two operationalizations are thus a single identification scheme applied to two outcome classes: the two-dimensional signal for real outcomes, where co-movement carries no circularity, and the technology signal alone for the return outcome, where it would.

A.5 Mechanism Chain in Complement Peers

Hypothesis 2 pins down the equilibrium response of peer market valuations. The mechanism through which the negative response on complement peers propagates into observable firm be-

havior unfolds across three real-economy margins: financial compression, labor co-optation, and innovation retreat.

A complement peer facing a credible permanent commoditization of its C-segment product experiences a structural reduction in pricing power. Revenue scaled by total assets contracts. At the same time, Alexy et al. (2018) observe that the release of an open-source public good by one firm provides a free, high-quality technological input to others, allowing them to reduce the internal expenditure required to produce equivalent capability. Complement peers adopting Ω as an input substitute cut their cost of goods sold at a rate that, on a scaled basis, can exceed the revenue contraction in the short run. Asset-scaled profitability therefore rises in the near term, as direct cost falls faster than revenue, even as the peer's competitive position deteriorates.

Hypothesis 3 (Financial Compression and Short-Run Profitability Expansion): *Complement peers exposed to a focal firm's open-source release exhibit declining sales and declining cost of goods sold, with the cost contraction outpacing the revenue decline so that asset-scaled profitability rises in the short run.*

The engineering labor force of a complement peer must maintain compatibility with the evolving open-source standard. In the platform-economics tradition of Gawer and Henderson (2007), standard-setting platforms redirect the focal effort of dependent developers toward the standard itself. In the open-source setting, this redirection is directly measurable as contribution activity on the focal firm's public repositories, separate from proprietary effort on the peer's own codebase.

Hypothesis 4 (Labor Co-optation): *Complement peer developers increase active contribution, measured through code commits, pull requests, and issues, to the focal firm's open-source repositories following the release.*

The reallocation of engineering effort onto the focal firm's commons reduces the proprietary capacity the complement peer can commit to frontier work. The declining marginal return to proprietary C-innovation, together with the continuing availability of Ω as an evolving baseline maintained by the focal firm, makes internal exploratory research less attractive than exploiting existing product lines for the complement peer. Capital reallocates away from exploratory R&D toward exploitation of existing product lines, and the peer's patent portfolio ages as new exploratory filings fall.

Hypothesis 5 (Innovation Retreat): *Complement peers reduce internal research and development expenditure and experience aging of their exploratory patent portfolios following a focal firm's open-source release.*

Hypotheses 3 through 5 together characterize the complement peer's adjustment path. The claim is not that all three must appear in every release but that they move together whenever the

release is economically consequential in the sense of the Kogan intensity screen described in Section 4.1. The baseline real-outcome specifications in Section 5 pool substitute and complement peers and therefore estimate an average treatment effect that, under Proposition 2, is driven predominantly by the complement-peer subsample. Section 6.2 decomposes the Sales and COGS responses along the same partition and verifies that the pooled estimates indeed load on the complement pool.

A.6 Welfare Considerations

The framework delivers signed predictions on producer surplus but leaves the consumer-welfare question open. On the producer side, the focal firm gains by Hypothesis 1, complement peers lose by Hypothesis 2, and substitute peers are approximately unchanged by Hypothesis 2. Across the three producer groups the focal-firm gain and the complement-peer loss carry the predicted opposite signs, with substitute peers approximately neutral; the paper documents this asymmetric repricing rather than attempting to compute a social-welfare net. On the consumer side, users of segment-C goods obtain a free substitute of quality V_Ω, which is weakly preferred to any priced alternative. Users of segment-A goods benefit from an expanded C-market that reduces the effective price of the $A + C$ bundle. Against these gains, the long-run consequence of standards lock-in through channel (iii) can reduce consumer variety by foreclosing entrants whose alternative interfaces would otherwise have emerged. The long-run sign on consumer surplus is therefore ambiguous. Consistent with the empirical stylebook of Seamans and Zhu (2014) and Wen and Zhu (2019), the paper positions its contribution as a documentation of the positive economic fact of the wealth transfer and leaves the welfare question open for downstream analysis.

B Appendix Robustness Discussion

B.1 Selection and Attenuation: Saturation Curve and Trimming

The headline stacked difference-in-differences specification weights all 1,671 release cohorts by their Kogan intensity $|\kappa_r|$. Two reviewer-relevant concerns merit a direct check. First, the set of included cohorts is fixed by mechanical rules, namely GitHub Archive coverage, a Compustat focal match, and the availability of a one-to-one propensity-score-matched control; a reviewer may worry that these rules implicitly select on outcome-relevant characteristics, particularly relative to a cutoff defined by cumulative repository popularity. Second, low-$|\kappa_r|$ releases could add noise that attenuates the weighted estimate. I address both with a saturation curve and a trimming-from-bottom exercise, reported in Table IA.10 and visualized in Figure IA3.

Saturation curve. Panel A of Table IA.10 ranks all cohorts by cohort-level $|\kappa_r|$ and re-estimates the headline $T \times Post$ coefficient on the Top-K subset for $K \in \{100, 250, 500, 1000, 1671\}$. The diagnostic is the following. If low-$|\kappa_r|$ releases carry attenuated but directionally consistent signal, adding them drags the equal-weighted estimate toward zero while leaving the $|\kappa_r|$-weighted estimate, which upweights the high-intensity releases, approximately unchanged; the two curves then coincide only on the highest-intensity subset and separate as K grows. The data show exactly this. For Sales/Assets the $|\kappa_r|$-weighted estimate is -0.065 at $K = 100$ and -0.060 at the full sample, a movement well inside one standard error, whereas the unweighted estimate drifts from -0.063 to -0.046, an attenuation of roughly 27%. R&D/Assets and ROA behave the same way: the weighted curve is approximately constant in K, while the unweighted curve attenuates away from it in absolute value as low-$|\kappa_r|$ releases are added. Two implications follow. The substantive effect is not driven by the cumulative-popularity Top 100 cutoff, since restricting to the highest-$|\kappa_r|$ Top 100 alone yields nearly the same coefficient as the full set of cohorts; and because the weighted estimate is stable across K, the full-sample weighted estimate is not diluted by the inclusion of low-intensity releases.

Trimming from the bottom. Panel B drops the bottom q% of cohorts by $|\kappa_r|$ and re-estimates the same specification for $q \in \{0, 25, 50, 75\}$. Under attenuation the coefficient should grow in absolute value as q rises. The weighted Sales coefficient is essentially flat across all four trims (-0.060, -0.060, -0.062, -0.065), while the unweighted Sales coefficient rises monotonically from -0.046 to -0.062, exactly the attenuation pattern that motivates the weighting choice. The two estimators meet at $q = 75$, where only the top $|\kappa_r|$ quartile of cohorts remains; the R&D and ROA panels show the same behavior. Low-$|\kappa_r|$ releases therefore contribute proportionally smaller, directionally consistent signal rather than dilutive noise, which is precisely the inverse-variance logic that justifies weighting by $|\kappa_r|$.

[Insert Appendix Table IA.10 Here]

B.2 Full-Sample VW CAR Robustness

The main-text market-reaction analysis in Section 6.3 uses the full sample of releases with return coverage. This appendix reports a complementary robustness check that varies the weighting scheme and the event window rather than the sample: Panel A of Appendix Table IA.12 reports both value-weighted and equal-weighted $\text{CAR}_{[-1,+1]}$ estimates for focal firms, substitute peers, and complement peers. The qualitative pattern in the main text, a positive focal CAR, a statistically insignificant substitute CAR, and a significantly negative complement CAR, is preserved under the equal-weighted and short-window variants. The focal, substitute, and complement dollar aggregates likewise preserve the direction and magnitude of the targeted wealth transfer documented in the main text.

[Insert Appendix Table IA.12 Here]

Appendix Figure IA4 shows the full-sample cumulative average abnormal return trajectory over the extended $[-10,+50]$ window as a visual complement.

[Insert Appendix Figure IA4 Here]

B.3 Core and Non-Core Segment Classification

The within-firm demand test reported in Table 10 compares the post-release sales path of a focal firm's Core business segments against its Non-Core segments. This subsection details how each Compustat business segment is assigned to the two groups and which focal firms the test can accommodate. Internet Appendix Table IA.6 reports the full segment-level mapping.

I classify a segment as Core if it is the focal firm's profit-center segment, segment A in the model, the segment whose demand the firm's strategic open-source releases are designed to expand by commoditizing the complementary segment; for these platform firms the profit center is typically the cloud, platform, or core product franchise. All other segments are Non-Core. The assignment is made once per segment from the firm's disclosure and held fixed over the panel. Where a firm reports historically renamed versions of the same line, such as Oracle's several cloud-license labels, I retain each named version as a separate segment because the Compustat extract carries them as distinct *snms* values. The difference-in-differences specification in Table 10 weights every firm-segment-year observation equally. This equalizes observations rather than firms, so a focal firm that reports more segments or more retained name variants does contribute more rows; the firm$\times$year and firm$\times$segment fixed effects absorb the resulting level differences, and identification

of the *IsCore* × *Post* coefficient comes only from within-firm-year Core-versus-Non-Core variation rather than from the number of retained variants.

Two focal firms report no segment that qualifies as Core: Apple's reported Compustat segment in the sample period is its Mobile Communication line, and Netflix reports only its streaming and legacy video-rental segments, all of which are Non-Core. With no Core segment, neither firm supplies the within-firm Core-versus-Non-Core contrast that identifies the coefficient on *IsCore* × *Post*, and I therefore exclude them from the demand panel. Internet Appendix Table IA.6 lists both the Core and Non-Core segment names for every focal firm.

Seven of the fifteen focal firms in the headline performance panel of Table 9 cannot enter the segment test. Amazon, Tesla, and Twilio have no business-segment data in the extract; Cisco reports a single Networking Systems segment; Salesforce reports two segments, Application Software and Enterprise Software, that describe the same product line and cannot be separated into Core and Non-Core; and Apple and Netflix report only the single non-core segment noted above. The within-firm demand analysis therefore uses the eight focal firms with a well-defined Core/Non-Core split: Adobe, Alphabet, IBM, Meta, Microsoft, Nvidia, Oracle, and Qualcomm.

[Insert Appendix Table IA.6 Here]

C Appendix Tables

Table IA.1. Selected Major Open-Source Releases by Focal Firms

This table lists twenty widely recognized open-source repositories drawn from the sample of roughly 1,700 release cohorts. *Release Date* is the cohort event date defined in Section 4.1, namely the repository's earliest open-source release recorded within the 2015–2024 sample window; for long-running projects this is a release during the window rather than the project's original public launch. Total Stars and Total Forks are the cumulative engagement metrics as of December 2024. The list is illustrative and is not used in estimation; the full event set enters the stacked difference-in-differences specification as described in Section 4.1.

Release Date	Company	Repository	Total Stars	Total Forks
2015-02-22	Meta	facebook/react	266,223	70,077
2015-11-09	Alphabet	tensorflow/tensorflow	222,055	113,229
2015-11-18	Microsoft	microsoft/vscode	219,500	60,918
2015-11-09	Alphabet	flutter/flutter	193,431	39,867
2015-03-26	Meta	facebook/react-native	142,423	34,803
2015-01-15	Alphabet	golang/go	140,768	24,778
2019-05-06	Microsoft	microsoft/PowerToys	122,816	8,839
2015-01-08	Alphabet	angular/angular	117,410	37,581
2015-08-10	Alphabet	kubernetes/kubernetes	113,262	47,065
2015-01-02	Microsoft	Microsoft/TypeScript	98,724	17,110
2016-02-05	Alphabet	tensorflow/models	88,698	57,029
2015-12-03	Apple	apple/swift	82,065	15,594
2018-01-22	Meta	facebook/create-react-app	72,792	30,901
2020-01-22	Microsoft	microsoft/playwright	69,346	4,500
2023-10-26	Microsoft	microsoft/generative-ai-for-beginners	65,001	33,310
2015-02-12	Alphabet	google/guava	56,084	15,081
2023-02-24	Meta	facebookresearch/llama	49,019	8,188
2019-06-23	Meta	facebook/docusaurus	47,124	10,857
2023-04-05	Meta	facebookresearch/segment-anything	47,319	5,795
2015-01-21	Meta	facebook/jest	43,326	8,433

Table IA.2. Face Validity of TNIC Peer Matching for Major Focal Firms

For each focal firm, this table lists its highest-similarity product-market peers from the lagged Hoberg-Phillips Text-based Network Industry Classification (TNIC) network used to define treated peers in Section 4.5. *TNIC* is the pairwise text-based product similarity between the focal firm and the peer, measured in the pre-release year. The parenthetical after each focal firm reports the total number of distinct TNIC peers it contributes to the panel; the firms listed are its five highest-similarity peers. The peers are the focal firms' recognizable product-market rivals, supporting the validity of the text-based treatment assignment.

Peer Firm	**TNIC**	**Peer Firm**	**TNIC**
Microsoft (401 peers)		**Alphabet** (113 peers)	
Nutanix Inc	0.110	Snap Inc	0.094
Citrix Systems Inc	0.109	Twitter Inc	0.083
Cisco Systems Inc	0.109	Pinterest Inc	0.061
Red Hat Inc	0.106	Adobe Inc	0.058
Adobe Inc	0.101	LinkedIn Corp	0.055
Meta (83 peers)		**Nvidia** (170 peers)	
Twitter Inc	0.133	Marvell Technology Inc	0.115
Pinterest Inc	0.096	Advanced Micro Devices	0.106
LinkedIn Corp	0.077	Qualcomm Inc	0.082
Snap Inc	0.068	Xilinx Inc	0.082
Zynga Inc	0.062	Sigma Designs Inc	0.080
Oracle (235 peers)		**Salesforce** (256 peers)	
New Relic Inc	0.129	New Relic Inc	0.130
Splunk Inc	0.113	Apigee Corp	0.121
Hewlett Packard Enterprise	0.106	QAD Inc	0.108
Cisco Systems Inc	0.105	Splunk Inc	0.108
Red Hat Inc	0.103	Zendesk Inc	0.107

Table IA.3. Matching Robustness: Full, PSM, CEM, and Entropy Balance

This table reports Kogan-weighted stacked Difference-in-Differences (DiD) estimates for the three main outcomes, Sales (Panel A), ROA (Panel B), and R&D (Panel C), under four alternative sample constructions. Column (1) uses the full unmatched sample. Column (2) uses the baseline Propensity Score Matching (PSM) sample. Column (3) applies Coarsened Exact Matching (CEM). Column (4) uses Entropy Balancing (EB). All dependent variables are normalized by total assets. All specifications include Cohort-Firm and Cohort-Time fixed effects, firm controls, and cluster standard errors at the firm level. t-statistics are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample period spans 2015–2024.

	(1)	(2)	(3)	(4)
Matching Method =	Full	PSM	CEM	EB
Panel A: Dependent Variable = Sales				
Treated × Post	-0.048**	-0.060***	-0.050***	-0.034
	(-2.48)	(-2.69)	(-3.16)	(-1.49)
Adjusted R^2	0.670	0.735	0.706	0.708
Mean of Dep. Var.	0.723	0.728	0.723	0.723
Observations	3,898,991	974,277	3,887,663	3,892,496
Panel B: Dependent Variable = ROA				
Treated × Post	0.058*	0.053**	0.065**	-0.005
	(1.94)	(2.00)	(2.43)	(-0.10)
Adjusted R^2	0.500	0.589	0.380	0.313
Mean of Dep. Var.	-0.095	-0.064	-0.089	-0.095
Observations	3,898,991	974,277	3,887,663	3,892,496
Panel C: Dependent Variable = R&D				
Treated × Post	-0.014***	-0.014***	-0.013***	-0.012***
	(-3.21)	(-2.77)	(-3.63)	(-3.26)
Adjusted R^2	0.786	0.788	0.695	0.701
Mean of Dep. Var.	0.093	0.107	0.092	0.093
Observations	3,900,467	974,385	3,888,670	3,893,505

Note: All models include Controls, Cohort-Firm FE, Cohort-Time FE, and Firm Clustering.

Table IA.4. Peer GitHub Activity Response: Unweighted Robustness

This table reports the unweighted versions of the Labor co-optation specification in main Table 5. Panel A uses the full sample of release cohorts; Panel B restricts to the Top 100 BigQuery-linked sub-sample. Both panels exclude the Kogan-intensity weights, so each cohort-firm-year observation enters the estimate equally rather than in proportion to its release intensity; releases with more peer-firm observations or fuller GitHub coverage still carry more rows. Outcomes are identical to those in Table 5: per-developer activity (total, active, and passive contributions) and the size of the active developer base. In the full-sample Panel A the unweighted estimates are generally larger than the Kogan-weighted main results; in the Top-100 Panel B, by contrast, the per-developer activity estimates are smaller, so dropping the weights does not uniformly strengthen the result. The passive-engagement effect, which is statistically indistinguishable from zero in the weighted main specification, becomes positive and significant in the unweighted full sample. All specifications include firm-level controls, cohort-firm and cohort-time fixed effects, and firm-clustered standard errors. t-statistics are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels.

Dependent Variable =	Activity per Developer			Developer Base
	(1) Any Work per Dev.	(2) Active Contrib.	(3) Passive Contrib.	(4) Ln(Working Devs + 1)
Panel A: Full sample				
Treated × Post	1.593***	1.502**	0.091***	0.245***
	(2.62)	(2.47)	(3.47)	(8.02)
Adjusted R^2	0.708	0.705	0.404	0.816
Mean of Dep. Var.	3.122	2.801	0.321	0.480
Observations	974,385	974,385	974,385	974,385
Panel B: Top 100 sub-sample				
Treated × Post	0.419***	0.378***	0.042*	0.171***
	(3.58)	(3.47)	(1.65)	(2.93)
Adjusted R^2	0.116	0.075	0.339	0.729
Mean of Dep. Var.	1.071	0.756	0.315	0.584
Observations	49,760	49,760	49,760	49,760
Controls	Yes	Yes	Yes	Yes
Cohort-Firm FE	Yes	Yes	Yes	Yes
Cohort-Time FE	Yes	Yes	Yes	Yes
Clustering	Firm	Firm	Firm	Firm
Weight	No	No	No	No

Table IA.5.

Long-Term Effect of Big Tech Open-Source Releases on Peer Firms' Patent Portfolio Age

This table presents the results of a stacked Difference-in-Differences (DiD) estimation of the effect of focal firms' open-source releases on the age of peer firms' patent portfolios, as a companion to the R&D investment estimates reported in main-text Table 6. The dependent variables are the average age of the overall patent portfolio (Columns 1 and 2) and the average age of the exploratory patent portfolio (Columns 3 and 4), measured in years. Columns (1) and (3) present a baseline OLS specification without fixed effects or controls. Columns (2) and (4) incorporate Cohort-Firm and Cohort-Time fixed effects along with firm-level controls, and weight the observations using the absolute value of the cohort-level Kogan (2017) innovation value. Control variables include Market Capitalization, Market-to-Book (M/B) Ratio, and Leverage. Standard errors are clustered at the firm level in the saturated specifications. t-statistics are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample period spans from 2015 to 2024.

	Patent Pf. Age		Exploratory Pf. Age	
	(1)	(2)	(3)	(4)
Treated × Post	0.476***	0.323***	0.744***	0.792***
	(21.11)	(2.84)	(27.86)	(5.19)
Market Cap		−0.074		0.771***
		(−0.51)		(2.99)
M/B Ratio		0.004		0.011
		(1.12)		(1.32)
Leverage		0.019		0.053
		(0.54)		(1.56)
Adjusted R^2	0.023	0.938	0.052	0.954
Cohort-Firm FE	No	Yes	No	Yes
Cohort-Time FE	No	Yes	No	Yes
Clustering	None	Firm	None	Firm
Weight	No	Yes	No	Yes
Mean of Dep. Var.	5.192	5.195	6.146	6.148
Observations	974,895	974,385	974,895	974,385

Table IA.6. Hand-Mapped Classification of Focal Firm Business Segments

This table reports the hand-coded classification of each focal firm's Compustat business segments into Core and Non-Core. A segment is classified as Core if it is the firm's profit-center segment (segment *A* in the model), the segment whose demand the firm's strategic open-source releases are designed to expand by commoditizing the complementary segment; for these platform firms the profit center is typically the cloud, platform, or core product franchise. All other reported segments are Non-Core. Both the Core and Non-Core segment names are listed so that the firm-segment observations identifying the $IsCore \times Post$ coefficient in Table 10 can be read off directly; the within-firm-year contrast is identified only for the eight firms that report at least one Core and one Non-Core segment. Apple and Netflix report no segment that qualifies as Core, so they contribute no within-firm Core-versus-Non-Core contrast and are excluded from the demand panel. Renamed variants of the same line appear as distinct *snms* values and are listed separately, as detailed in Appendix Section B.3. The sample period spans 2010–2023.

Focal Firm	Core Business Segments (10-K *snms*)	Non-Core Business Segments (10-K *snms*)
Adobe	Digital Media; Digital Experience; Digital Marketing	Creative Solutions; Creative & Interactive Solutions; Digital Media Solutions; Omniture; Knowledge Worker; Enterprise; Platform; Print and Publishing; Publishing; Publishing & Advertising
Alphabet	Google; Google Cloud; Google Services; Internet Software & Services; Web Search Portal	Other Bets; Motorola Mobile; Mobile; Corporate
Apple	none classified as Core	Mobile Communication
IBM	Software; Cognitive Solutions; Technology Services & Cloud; Cloud & Cognitive Software	Global Business Services; Global Technology Services; Global Financing; Financing; Consulting; Infrastructure; Systems; Systems & Technology; Systems & Technology Group; Systems Hardware; Systems and Technology Group; Other and Adjustments
Meta	Family of Apps; Social Networking	Reality Labs
Microsoft	Intelligent Cloud; Server and Tools; Productivity and Business; Microsoft Business Division; Commercial – Licensing; Commercial – Other	More Personal Computing; Windows Division; Windows & Windows Live; Windows & Windows Live Division; Online Services Division; Online Services Business Division; Entertainment & Devices; Entertainment & Devices Division; D & C – Computing & Gaming; D & C – Licensing; D & C – Phone Hardware; D & C – Other; Corporate & Other; Unallocated & Other
Netflix	none classified as Core	Streaming Entertainment Serv; Video Tape Rental
Nvidia	Compute & Networking; GPU; Tegra Processor	Graphics; CPB; MCP; PSB; All Other
Oracle	Cloud Infrastructure; Cloud SaaS & PaaS; Cloud Services; Cloud and License; Cloud & On-Premise Software; New Software Licenses; New Software Licenses & Cloud Software Subscriptions; Software License Updates; Software License Updates & Product Support	Hardware; Hardware Products; Hardware Support; Hardware Systems Products; Hardware Systems Support; Services; Services Business; Consulting; Education; On Demand
Qualcomm	QCT (Qualcomm CDMA Technologies, the semiconductor product segment)	QTL; QSI; QWI; Reconciling Items

Table IA.7. Illustrative Examples of Substitute vs. Complement Peers (2D Classification)
This table illustrates how the two-dimensional classifier assigns peer firms into *Substitute* vs. *Complement* cohorts for three open-source releases by focal technology firms. A peer is labeled a *Substitute* if its CPC patent cosine similarity with the focal firm is *above* the cohort median (high technology overlap, a direct technological rival) and its pre-event $[t-250, t-30]$ industry-adjusted return correlation with the focal firm is *negative*; a peer is labeled a *Complement* if its CPC similarity is *below* the cohort median (an adjacent technology layer) and the pre-event correlation is *positive*. Peers for which the two signals disagree are dropped as ambiguous. Firms appearing in multiple panels are excluded to showcase distinct cases. *CPC Sim.* is cosine similarity (0–1) between the peer's and the focal firm's patent-CPC profiles, classified relative to the cohort median; *Corr* is the peer's pre-event $[t-250, t-30]$ industry-adjusted return correlation with the focal firm. High-overlap firms predominantly co-move with the focal firm, so the substitute bucket (high overlap with negative co-movement) sits near the sign boundary; the partition's validity rests on the regression estimates of Table 7, not on individual borderline pairs. Sample period: 2015–2024.

Substitute Peers (CPC > med. & Corr < 0)			**Complement Peers (CPC < med. & Corr > 0)**		
Peer Firm	CPC Sim.	Corr	Peer Firm	CPC Sim.	Corr
Panel A: microsoft/vscode			(Focal: Microsoft Corp.) — *developer IDE*		
Red Hat Inc.	0.97	−0.07	Cisco Systems Inc.	0.51	+0.37
Appian Corp.	0.95	−0.09	F5 Inc.	0.52	+0.24
ServiceSource Intl Inc.	0.91	−0.13	Qualcomm Inc.	0.34	+0.21
Panel B: adobe/react-spectrum			(Focal: Adobe Inc.) — *UI component library*		
PDF Solutions Inc.	0.98	−0.12	Roku Inc.	0.07	+0.28
Blackbaud Inc.	0.31	−0.16	Zscaler Inc.	0.09	+0.28
Cloudera Inc.	0.31	−0.24	Dolby Laboratories Inc.	0.02	+0.27
Panel C: microsoft/react-native-windows (Focal: Microsoft Corp.) — *cross-platform app framework*					
Splunk Inc.	0.98	−0.12	Akamai Technologies Inc.	0.50	+0.14
Barracuda Networks Inc.	0.95	−0.11	CDK Global Inc.	0.31	+0.17
FalconStor Software Inc.	0.94	−0.06	Flex Ltd.	0.02	+0.26

Table IA.8.

Robustness: Substitute vs. Complement Peers under One-Dimensional CPC Classification

This table re-estimates the main Substitute-versus-Complement decomposition of Table 7 under a simpler one-dimensional classification that uses technology overlap alone and retains every peer, rather than the two-dimensional intersection of the main specification that drops ambiguous off-diagonal peers. A peer is labeled a Substitute if its CPC cosine similarity with the releasing firm is above the cohort median and a Complement if it is below; return co-movement is not used. The dependent variables Sales and COGS are normalized by total assets. All specifications include Cohort-Firm and Cohort-Time fixed effects and firm-level controls, and observations are weighted by the absolute value of the cohort-level Kogan (2017) innovation value. Standard errors are clustered at the firm level. t-statistics are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample period spans 2015–2024.

Dependent Variable =	Sales		COGS	
Peer Firm Type =	(1) Substitute	(2) Complement	(3) Substitute	(4) Complement
Treated × Post	-0.0335	-0.0928***	-0.0376**	-0.0643***
	(-1.36)	(-3.20)	(-1.97)	(-2.79)
Adjusted R^2	0.654	0.725	0.661	0.747
Controls	Yes	Yes	Yes	Yes
Cohort-Firm FE	Yes	Yes	Yes	Yes
Cohort-Time FE	Yes	Yes	Yes	Yes
Clustering	Firm	Firm	Firm	Firm
Weight	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.666	0.729	0.288	0.352
Observations	554,830	559,006	554,830	559,006

Table IA.9. Heterogeneous Effects on
Substitute vs. Complement Peers under a Large-Language-Model Economic Classification

This table re-estimates the Substitute–Complement heterogeneity of Table 7 under an alternative classification: each peer is labeled a Substitute (competes in the focal firm's profit center) or a Complement (operates in the layer the release commoditizes) by majority vote of three large language models (Claude Opus, Sonnet, and Haiku) reading its core business, independent of patent overlap and returns and blind to all outcomes. Each column is a separate stacked DiD for the dependent variable in its header; the specification is identical to Table 7 ($|\kappa|$-weighted, Cohort-Firm and Cohort-Time fixed effects, firm-level controls, firm-clustered standard errors), and Sales, COGS, and R&D are scaled by total assets. Panel A reports Substitute peers, Panel B Complement peers. The unanimous-vote subset sharpens the contrast further, with the complement-peer ROA coefficient rising to 0.118. t-statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable =	Sales	COGS	ROA	R&D
	(1)	(2)	(3)	(4)
Panel A: Substitute Peers				
Treated × Post	-0.0285	-0.0311*	0.0356	-0.0072
	(-1.03)	(-1.82)	(0.89)	(-0.58)
Adjusted R^2	0.620	0.617	0.706	0.864
Mean of Dep. Var.	0.684	0.323	-0.065	0.087
Observations	386,076	386,076	386,076	386,172
Panel B: Complement Peers				
Treated × Post	-0.1162***	-0.0571**	0.0840***	-0.0297***
	(-2.98)	(-2.11)	(2.78)	(-3.33)
Adjusted R^2	0.622	0.622	0.708	0.864
Mean of Dep. Var.	0.673	0.310	-0.065	0.087
Observations	391,475	391,475	391,475	391,574
Controls	Yes	Yes	Yes	Yes
Cohort-Firm FE	Yes	Yes	Yes	Yes
Cohort-Time FE	Yes	Yes	Yes	Yes
Clustering	Firm	Firm	Firm	Firm
Weight	Yes	Yes	Yes	Yes

Table IA.10. Selection and Attenuation: Saturation Curve and Trimming-from-Bottom

Panel A re-estimates the headline weighted stacked DiD $T \times Post$ coefficient on Top-K subsets of release cohorts ranked by Kogan intensity $|\kappa_r|$ (descending), for K from 100 up to the full sample of 1,671 cohorts (the rightmost block is labeled Top-1700 for readability). Panel B reports the same coefficient with the bottom $q\%$ of cohorts by $|\kappa_r|$ dropped. "Wgt" is the $|\kappa_r|$-weighted specification; "Unw" is unweighted. All specifications include Cohort-Firm and Cohort-Time fixed effects and firm-level controls; standard errors are clustered by firm and t-statistics are reported in parentheses. The sample is the full stacked DiD panel. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Saturation curve, Top-K cohorts by $|\kappa|$

K	Top-100		Top-250		Top-500		Top-1000		Top-1700	
	Wgt	Unw	Wgt	Unw	Wgt	Unw	Wgt	Unw	Wgt	Unw
Sales/Assets	−0.0649**	−0.0633**	−0.0667***	−0.0657**	−0.0642***	−0.0608***	−0.0606***	−0.0504**	−0.0597***	−0.0463**
	(-2.56)	(-2.49)	(-2.59)	(-2.52)	(-2.69)	(-2.71)	(-2.69)	(-2.39)	(-2.69)	(-2.39)
R&D/Assets	−0.0089	−0.0092	−0.0129**	−0.0145**	−0.0139**	−0.0157***	−0.0141***	−0.0150***	−0.0138***	−0.0125***
	(-1.45)	(-1.53)	(-2.19)	(-2.47)	(-2.56)	(-3.07)	(-2.76)	(-3.04)	(-2.77)	(-2.78)
ROA	0.0378	0.0427	0.0525*	0.0600*	0.0544*	0.0603**	0.0537**	0.0549**	0.0528**	0.0474**
	(1.36)	(1.48)	(1.73)	(1.86)	(1.91)	(2.17)	(1.98)	(2.01)	(2.00)	(2.02)
Cohorts (events)	100	100	250	250	500	500	1000	1000	1671	1671
Obs.	69,055	69,055	179,122	179,122	336,798	336,798	638,158	638,158	974,277	974,277

Panel B: Trimming-from-bottom, drop bottom q% of cohorts by $|\kappa|$

Bottom dropped	$q = 0\%$		$q = 25\%$		$q = 50\%$		$q = 75\%$	
	Wgt	Unw	Wgt	Unw	Wgt	Unw	Wgt	Unw
Sales/Assets	−0.0597***	−0.0463**	−0.0600***	−0.0487**	−0.0616***	−0.0536**	−0.0650***	−0.0623***
	(-2.69)	(-2.39)	(-2.69)	(-2.37)	(-2.70)	(-2.50)	(-2.68)	(-2.68)
R&D/Assets	−0.0138***	−0.0125***	−0.0139***	−0.0136***	−0.0142***	−0.0156***	−0.0139**	−0.0160***
	(-2.77)	(-2.78)	(-2.76)	(-2.85)	(-2.74)	(-3.14)	(-2.51)	(-3.01)
ROA	0.0528**	0.0474**	0.0532**	0.0519**	0.0538**	0.0560**	0.0539*	0.0601**
	(2.00)	(2.02)	(2.00)	(2.02)	(1.96)	(2.04)	(1.86)	(2.06)
Cohorts (events)	1671	1671	1253	1253	835	835	417	417
Obs.	974,277	974,277	762,804	762,804	547,253	547,253	286,568	286,568

Wgt = Kogan $|\kappa|$-weighted; Unw = unweighted. t-stats in parentheses, clustered by gvkey.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table IA.11. Weighting Robustness: Continuous vs. Rank-Based Kogan Weights

This table addresses the concern that using the continuous absolute Kogan (2017) innovation value as a weight could be driven by a small number of extreme cohorts. The columns compare the baseline continuous-weight specification to an alternative that replaces the weight with the within-cohort rank (quartile rank, rescaled). The dependent variables are Sales, Return on Assets (ROA), and R&D expenditure, all normalized by total assets. Coefficients are robust across the three main outcomes, confirming that the main result is not driven by extreme-intensity observations. All specifications include Cohort-Firm and Cohort-Time fixed effects, firm-level controls, and cluster standard errors at the firm level. t-statistics are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample period spans 2015–2024.

Dependent Variable =	Sales		ROA		R&D	
Weighting Scheme =	(1) Continuous	(2) Rank-Based	(3) Continuous	(4) Rank-Based	(5) Continuous	(6) Rank-Based
Treated × Post	-0.060***	-0.050**	0.053**	0.052**	-0.014***	-0.014***
	(-2.69)	(-2.48)	(2.00)	(2.07)	(-2.77)	(-2.99)
Adjusted R^2	0.735	0.739	0.589	0.476	0.788	0.745
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Cohort-Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Cohort-Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Clustering	Firm	Firm	Firm	Firm	Firm	Firm
Mean of Dep. Var.	0.728	0.728	-0.064	-0.064	0.107	0.107
Observations	974,277	974,277	974,277	974,277	974,385	974,385

Table IA.12. Full-Sample Robustness: EW vs. VW $CAR_{[-1,+1]}$ by Peer Type

This table reports equal-weighted (EW) and value-weighted (VW) cumulative abnormal returns over the $[-1,+1]$ event window, along with aggregate dollar-value impact, for the full sample of releases (i.e., every release for which peer-firm return data are available, without the top-100 restriction imposed in the main text). Abnormal returns are estimated from the FF4 model. Firms are weighted by pre-event market capitalization in the VW specification. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels.

Group	**N**	**EW CAR%**	**EW *t***	**VW CAR%**	**VW *t***	**Agg. $B**
Focal Firms	1,707	$+0.22$	$(+2.36)^{**}$	$+0.17$	$(+3.85)^{***}$	$+1,905$
Substitute Peers	31,041	$+0.08$	$(+1.54)$	$+0.06$	$(+1.42)$	$+1,055$
Complement Peers	17,402	-0.13	$(-2.88)^{***}$	-0.13	$(-3.12)^{***}$	-143

D Appendix Figures

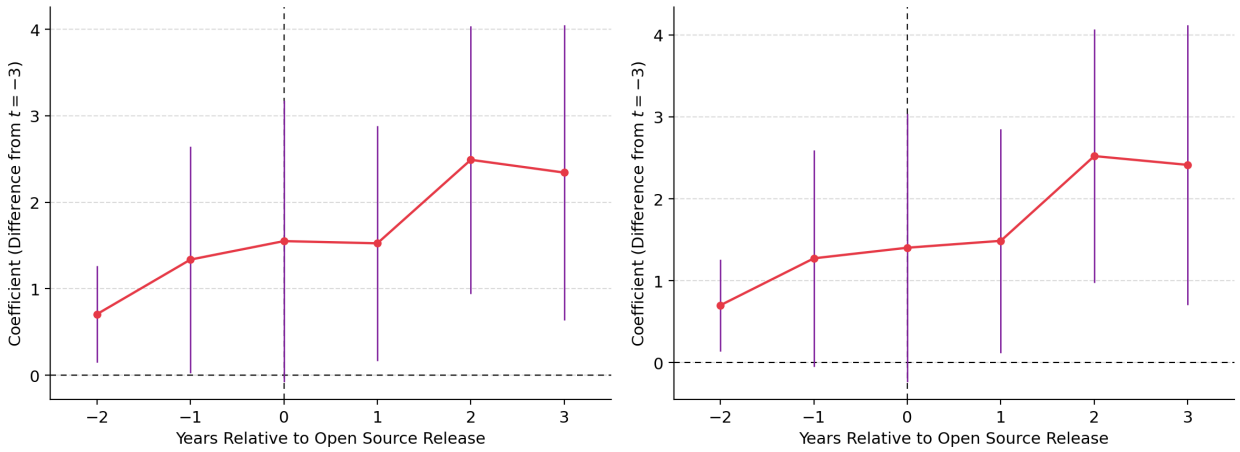


Figure IA1. Dynamic Effect on Peer Firm GitHub Activity per Developer (Full Sample)
This figure plots the dynamic treatment effect of focal firms' open-source releases on peer firms' per-developer GitHub activity using the full sample of approximately 1,700 cohorts, matching Panel A of Table 5. Panel (A) shows any contribution per working developer; Panel (B) shows active contributions per working developer. This figure is the dynamic event-study counterpart of the full-sample static estimate reported in main-text Table 5. As discussed in Section 5.3, the broader sample exhibits a positive pre-treatment trajectory consistent with the secular GitHub-adoption ramp affecting most peer firms during 2014–2017; the pooled DiD estimate in Table 5 Panel A controls for this trend through the cohort $\times$ year fixed effects, which is why the static specification rather than this dynamic one is used as the headline. Coefficients are plotted relative to $t = -3$ with 95% confidence intervals.

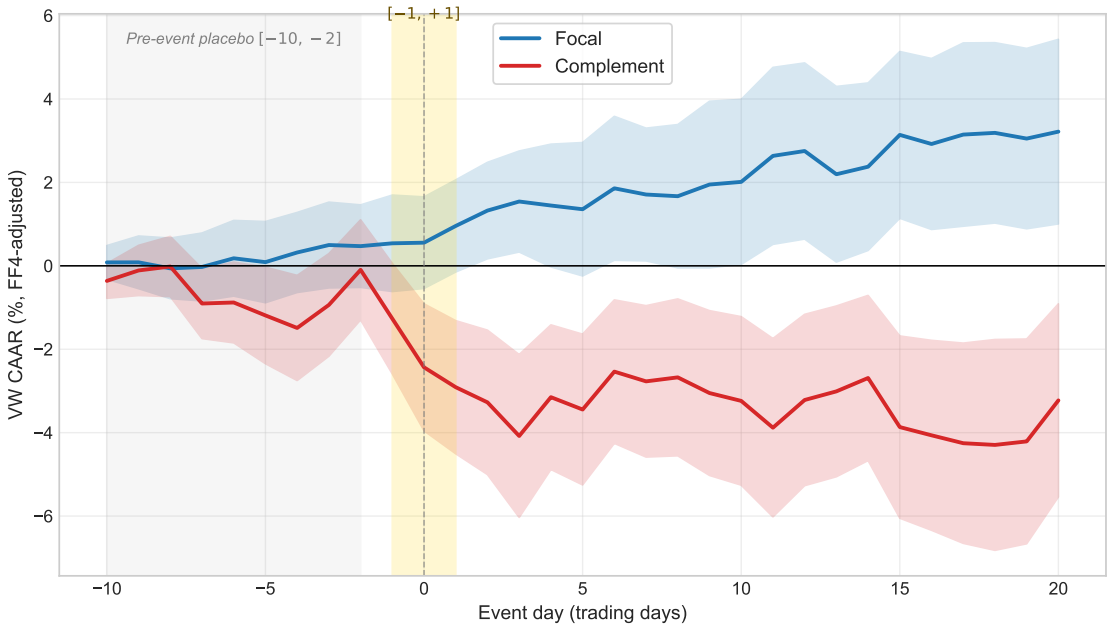


Figure IA2. CAAR Dynamics for the Top 100 Strategic Releases: Focal vs. Complement Peers

This figure traces the value-weighted cumulative average abnormal return for focal firms and complement peers from trading day -10 through day $+20$ relative to the release day. The sample is the top 100 strategically influential open-source releases. Abnormal returns are estimated from the Fama-French four-factor (FF4) model and weighted by pre-event market capitalization. Shaded bands denote 95% confidence intervals constructed from the effective-sample-size weighted cross-sectional variance. The gray-shaded region marks the pre-event placebo window $[-10, -2]$; both groups exhibit flat pre-trends. The gold-shaded region marks the baseline event window $[-1, +1]$. Substitute peers are omitted for clarity because their reaction is small and second-order relative to the focal and complement responses.

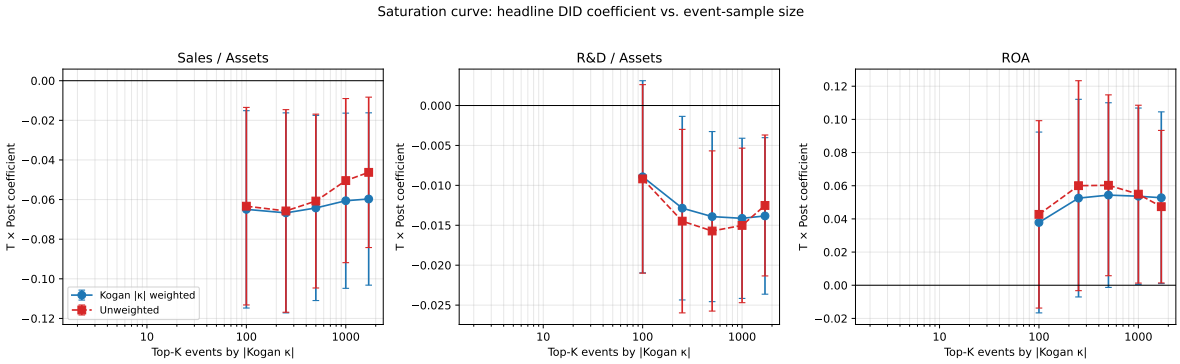


Figure IA3. Saturation Curve: Headline DiD Coefficient vs. Event-Sample Size

Each panel plots the weighted stacked-DiD $T \times Post$ coefficient with its 95% confidence interval against the Top-K subset of release cohorts ranked by Kogan intensity $|\kappa_r|$, for the three main outcomes (Sales/Assets, R&D/Assets, ROA). The solid line is the $|\kappa_r|$-weighted estimate; the dashed line is the unweighted estimate. The two estimates coincide on the highest-intensity subset, and as lower-$|\kappa_r|$ cohorts are added toward the full sample of approximately 1,700 the unweighted estimate attenuates toward zero while the weighted estimate stays put, indicating that low-$|\kappa_r|$ releases contribute attenuated but directionally consistent signal that dilutes the equal-weighted estimate. Companion to Table IA.10; the sample is the full stacked DiD panel.

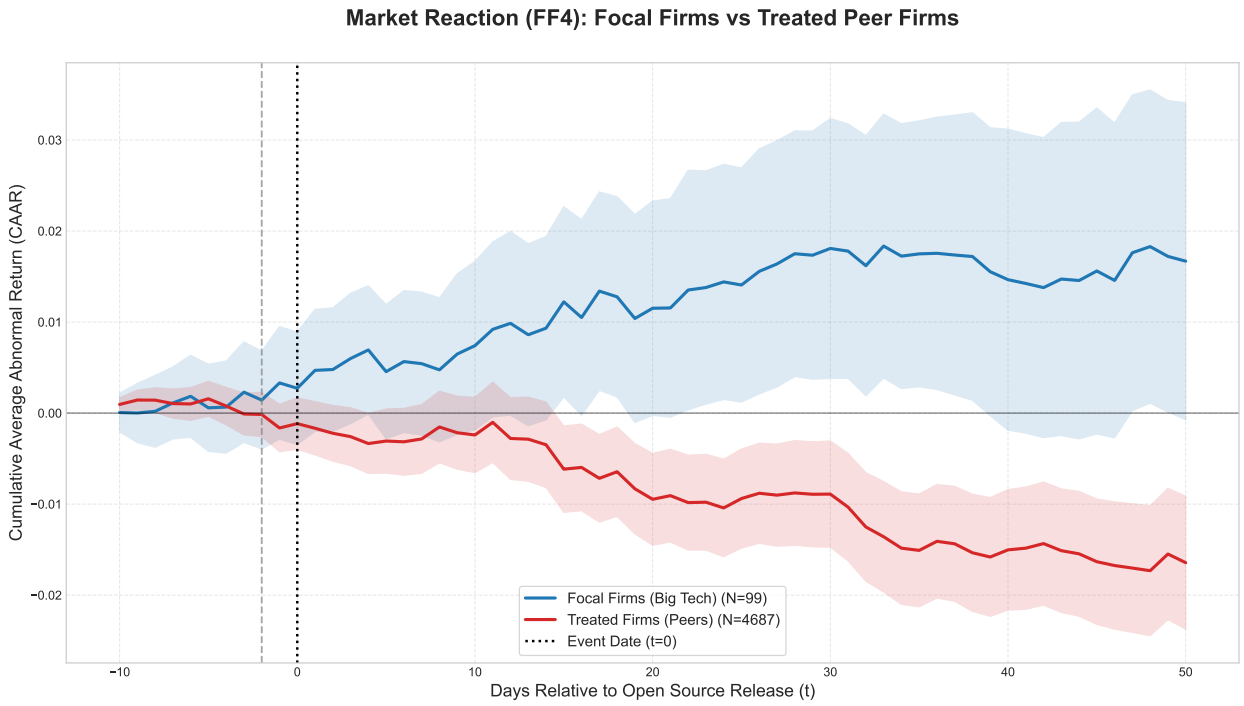


Figure IA4. Full-Sample CAAR: Focal vs. Treated Peers over $[-10, +50]$

Cumulative average abnormal returns (CAAR) for focal tech giants and the pooled treated peer group over the trading-day window $[-10, +50]$ relative to the release day (Day 0), estimated from the FF4 model with 150-day estimation windows ending 30 days prior to the event. Shaded bands represent 90% confidence intervals. Unlike Figure IA2, which uses the top 100 sample and shows focal and complement separately, this version uses the full sample and pools substitute and complement peers; the pooled peer trajectory lies between the two individual partitions.