# Commoditize Your Complement:
# The Cannibalization of Peer Firms by Open-Source Shocks

Peter Young Won *

March 2026

**ABSTRACT**

This paper investigates how dominant technology firms release free open-source software to commoditize complementary markets. By releasing free open-source software, these firms effectively devalue competitors' proprietary assets. Exploiting the staggered release of major GitHub repositories, I find that exposed rivals experience a 7.6% decline in market valuation and a 5.2% reduction in sales. To absorb this revenue compression, competitors adopt the new open-source software more actively to reduce costs, thereby inflating short-term profit margins. However, this defensive restructuring alters long-term resource allocation: direct rivals decrease R&D expenditures by 14.8%, curtail exploratory innovation, dedicate 29.3% more human capital to engaging with the focal firm's open-source projects, and exhibit decreased product market fluidity. These effects are concentrated among more direct competitors, illustrating how the strategic distribution of open-source software structurally reorganizes industry competition.

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

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

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

# Disclosure Statement

**Research Support:** This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

**Conflict of Interest:** The author, Peter Y. Won, declares no financial relationships or other potential conflicts of interest that could inappropriately influence or bias the findings, interpretations, or conclusions of this study.

**Data Access and Responsibility:** As the sole author, I had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

# 1 Introduction

Why do leading technology firms invest massive capital to develop advanced software, only to release it to the public for free? Over the past two decades, open-source software (OSS) releases by dominant platforms like Alphabet, Meta, and Microsoft have become the foundational infrastructure of the modern digital economy. While this phenomenon is widely celebrated as a collaborative public good, it raises a fundamental economic question: is the proliferation of free software truly a benign technological spillover, or does it mask a deliberate corporate strategy? Theoretical frameworks suggest that dominant platforms can strategically distribute free software to commoditize complementary technological layers, thereby stimulating demand for their highly profitable core businesses (Parker and Van Alstyne, 2005; Gawer and Henderson, 2007). Yet, the extent to which this "commoditize your complement" strategy actually cannibalizes the pricing power and proprietary resource advantages of incumbent rivals remains a critical, unresolved question.

Evaluating this mechanism is empirically challenging. Researchers rarely observe the precise product-market exposure of incumbent firms to a new open-source technology, nor can they easily track the subsequent reallocation of engineering human capital. Consequently, empirical literature has largely treated open-source releases as benign public goods that boost aggregate productivity (Nagle, 2019; Hoffmann et al., 2024), leaving their predatory impact on direct proprietary rivals as an unexplored black box. To overcome these data limitations, I construct a novel, multi-dimensional dataset that links the top 100 most influential open-source repositories released by big tech firms on GitHub (2015–2024) to the financial and innovative outcomes of U.S. public firms. Using text-based network industry classifications (TNIC) (Hoberg et al., 2014), I identify the exact product market rivals exposed to these technological shocks. I then exploit the staggered timing of these releases in a stacked difference-in-differences framework to test how open-source entry reshapes competitors' cost structures, innovation, and labor allocation.

My primary empirical results demonstrate that big tech open-source releases operate as a strategic weapon that structurally cannibalizes proprietary rivals. First, I find that the stock market immediately prices in this predatory threat: exposed rivals suffer a 7.6 percent annualized decline in market valuation following the focal platform's OSS release. This wealth destruction is fundamentally driven by top-line cannibalization, evidenced by a structural 5.2 percent decline in the rivals'

sales. Faced with this severe revenue compression, incumbent rivals are forced into a survival-driven operational restructuring. To avoid immediate financial distress, they aggressively slash their direct manufacturing costs (COGS) and organizational overhead (SG&A) by strategically free-riding on the newly available open-source infrastructure (Emery et al., 2025). Paradoxically, because these cost reductions temporarily outpace the initial decline in sales, treated firms experience a mechanical inflation of their short-term profit margins (ROA and EBITDA).

Second, while this defensive cost-cutting creates a temporary illusion of financial stability, I document a devastating long-term strategic toll on independent innovation and human capital. By substituting proprietary development with external open-source public goods, treated firms decrease their internal R&D expenditures by 14.8 percent. Consequently, the average age of their patent portfolios increases significantly, and their share of exploratory, frontier-level patents plummets. Crucially, I uncover a massive, structural reallocation of engineering labor. Utilizing real-time developer engagement logs from GitHub, I find that active development staffing at rival firms dedicated to the focal platform's open-source project increases by nearly 30 percent, with active coding contributions surging by over 50 percent. Dominant platforms essentially co-opt the highly skilled proprietary talent of their competitors, forcing rivals to finance the human capital required to sustain the platform's open ecosystem (Gupta et al., 2024).

Finally, this extraction of internal R&D and engineering labor culminates in a severe chilling effect on the rivals' actual product offerings, evidenced by a significant decrease in product market fluidity. Crucially, heterogeneity analysis reveals the dual nature of these releases. For peripheral firms (Low Threat), the open-source release acts as a benign technological input that allows them to efficiently expand their product scope. However, for direct product market rivals (High Threat), the exact same release functions as a devastating predatory weapon. The extreme hollowing out of R&D, the massive surrender of engineering labor, and the complete paralysis of product market scope are strictly concentrated among these direct competitors.

Underpinning these findings is a novel, multi-dimensional dataset that bridges real-time developer activity with corporate financial and product market networks. Specifically, I integrate high-frequency developer engagement logs from the top 100 most influential open-source repositories released by big tech firms on GitHub (2015–2024) with fundamental accounting data from Compustat, market valuations from CRSP, and intellectual property data from the USPTO. To precisely

identify which incumbent firms are exposed to these open-source shocks, I employ the text-based network industry classifications (TNIC) developed by Hoberg et al. (2014); Hoberg and Phillips (2016). This granular, text-based product space mapping allows me to construct tightly matched counterfactuals and leverage the staggered timing of major open-source releases in a robust stacked difference-in-differences framework (Baker et al., 2022), effectively isolating the causal impact of platform-driven commoditization from pre-existing industry trends.

The contributions of this paper are threefold. First, this study provides the first large-scale empirical quantification of the "commoditize your complement" strategy (Parker and Van Alstyne, 2005). While prior theoretical work has elegantly modeled the incentives for platforms to subsidize complementary markets (Economides and Katsamakas, 2006), recent empirical literature has primarily focused on the positive externalities of open-source software, such as aggregate productivity gains (Nagle, 2019), startup funding (Conti et al., 2025), and general firm value enhancements (Hoffmann et al., 2024; Zhao, 2025). This paper offers the empirical bridge to the unexplored, darker side of platform economics. By utilizing high-frequency GitHub data integrated with text-based product networks, I identify information-driven commoditization as a fundamental mechanism of platform dominance, demonstrating how this strategy operates in real-time to destroy the expected future cash flows and top-line revenues of proprietary competitors.

Second, this paper significantly extends the strategic management and corporate finance literature on the Resource-Based View (RBV) and open innovation (Alexy et al., 2018; Mkrtchyan et al., 2025). Recent studies highlight how firms transition to open-source software to optimize input costs (Haese and Peukert, 2025) and capture private value from open innovation (Emery et al., 2025; Coleman et al., 2025). However, these studies primarily examine open innovation from the perspective of the focal firm or the willing adopter gaining ecosystem advantage. This paper shifts the paradigm by documenting the involuntary surrender of resources by competitor firms. I show that when a platform establishes a new open standard, it not only degrades the value of competitors' proprietary assets but literally co-opts their human capital (Gupta et al., 2024). This provides novel evidence that in digital markets, labor reallocation and innovation decay are the direct, hidden costs of defensive technological free-riding.

Third, these findings offer critical insights for modern antitrust enforcement and digital policy. Traditional regulatory frameworks predominantly evaluate anti-competitive behavior through the

lens of predatory pricing, output restriction, or explicit exclusion in core markets. However, as data and digital infrastructure increasingly function as networked assets (Bian et al., 2025), this study demonstrates that dominant digital platforms can exercise profound market power through the predatory free distribution of complements. By giving away world-class software, tech giants evade traditional antitrust scrutiny while systematically starving rivals of revenue, stifling independent exploratory innovation, and monopolizing the industry's engineering talent. This paper bridges economic measurement with policy, highlighting that the true cost of "free" open-source software is the subtle but systematic neutralization of the competitive ecosystem.

The remainder of the paper is organized as follows. Section 2 outlines the theoretical framework. Section 3 describes the data construction. Section 3.5 details the stacked difference-in-differences empirical strategy. Section 4 presents the baseline empirical findings regarding market reaction, revenue cannibalization, cost-cutting, innovation decay, and labor co-optation. Section 5 provides the heterogeneity analysis. Section 6 concludes.

# 2 Theoretical Framework and Hypothesis Development

To understand the profound structural shifts documented in this study, it is necessary to formally model the strategic interactions between a dominant technology platform (the focal firm) and its proprietary software competitors (incumbent rivals). In this section, I integrate the economics of two-sided network effects (Parker and Van Alstyne, 2005) with the resource-based view (RBV) of the firm (Alexy et al., 2018) to develop a unified theoretical framework. This model explains both the focal platform's incentive to commoditize its complements and the rival firms' subsequent defensive adjustments, leading to a set of testable empirical hypotheses.

## 2.1 Economic Intuition: Two-Sided Networks and Resource Commoditization

The theoretical foundation of this study rests on two interrelated economic mechanisms: the strategic pricing of information goods in two-sided markets, and the subsequent devaluation of proprietary resources.

First, to understand why a profit-maximizing tech giant would invest massive capital into a high-quality software project only to release it for free, we must view the firm as a multi-product platform. Such dominant platforms typically operate in two distinct but highly interdependent markets: a highly profitable core market, such as cloud infrastructure or hardware, and a complementary software market. The defining characteristic of this ecosystem is the presence of strong cross-network externalities. As demonstrated by Parker and Van Alstyne (2005), when the widespread adoption of complementary software significantly stimulates the demand for the core infrastructure, the platform benefits from a massive cross-market subsidy effect. Because giving away the software for free maximizes user adoption and exponentially drives traffic to the profitable core business, this subsidy effect easily dominates the direct costs of software development. Economically, the platform's optimal strategy is to intentionally commoditize the complementary technological layer into perpetuity.

Second, the introduction of this zero-priced, high-quality open-source infrastructure fundamentally disrupts the competitive equilibrium for incumbent rivals. Traditional strategy literature, particularly the Resource-Based View (RBV), posits that a firm sustains its competitive advantage and generates monopoly rents by holding proprietary resources that are valuable, rare, and inimitable (VRI) (Barney, 1991). Before the open-source release, incumbent rivals thrived by closely guarding their proprietary algorithms and software architectures.

However, Alexy et al. (2018) theorize that the strategic release of open-source technology transforms a previously proprietary asset class into a ubiquitous common-pool resource. When the focal platform establishes its open-source project as the new, free industry standard, it instantaneously destroys the rarity and inimitability of the rivals' proprietary code. The rival's core asset transitions from a rent-generating VRI resource into a rapidly depreciating legacy system, completely stripping away its pricing power.

By integrating these two theories, a clear predatory mechanism emerges. The focal platform intentionally leverages cross-network externalities to dump high-quality public goods into the market, securing its core monopoly while systematically dismantling the proprietary resource advantages of incumbent rivals. The following subsections formally model this dynamic, developing specific hypotheses regarding how this shock sequentially triggers wealth destruction, cost restructuring, innovation decay, and labor co-optation.

## 2.2 The Economics of Open-Source Commoditization

Consider a dominant platform firm ($F$) that operates in two interconnected markets: a core infrastructure market ($C$), such as cloud computing or hardware, and a complementary software market ($OSS$). The platform's total profit function is given by:

$$\Pi_F = P_C \cdot D_C(P_C, D_{OSS}) + P_{OSS} \cdot D_{OSS}(P_{OSS}, D_C) - C(D_C, D_{OSS}) \tag{1}$$

where $P$ denotes price, $D$ denotes demand, and $C(\cdot)$ is the joint cost function. A critical feature of digital platforms is the presence of strong cross-network externalities. An expansion in the complementary software ecosystem significantly increases the demand for the core infrastructure ($\frac{\partial D_C}{\partial D_{OSS}} > 0$).

Maximizing the platform's profit with respect to the price of the complementary software ($P_{OSS}$) yields the first-order condition:

$$\frac{\partial \Pi_F}{\partial P_{OSS}} = D_{OSS} + \left( P_{OSS} - \frac{\partial C}{\partial D_{OSS}} \right) \frac{\partial D_{OSS}}{\partial P_{OSS}} + P_C \frac{\partial D_C}{\partial D_{OSS}} \frac{\partial D_{OSS}}{\partial P_{OSS}} = 0 \tag{2}$$

Following Parker and Van Alstyne (2005), when the cross-market subsidy effect (the final term) is sufficiently large, the optimal strategy for the dominant platform is to set $P^*_{OSS} = 0$ into perpetuity. By releasing high-quality open-source software, the platform effectively commoditizes the complementary market to stimulate demand for its highly profitable core business.

## 2.3 Value Destruction and Defensive Free-Riding

When the focal platform releases open-source software ($\Omega$) at a price of zero, it fundamentally disrupts the market for incumbent rivals ($R$) selling proprietary software. Consumers evaluate the utility of the open-source alternative ($U_{OSS} = V_{OSS} - 0$) against the proprietary software ($U_R = V_R - P_R - \tau$), where $V$ represents product value and $\tau$ denotes switching costs (Zhu and Zhou, 2012).

To prevent customer defection ($U_R \geq U_{OSS}$), the rival firm is forced to severely compress its proprietary pricing: $P^*_R \leq (V_R - V_{OSS}) - \tau$. As the quality of the free open-source software increases, the rival's pricing power collapses, directly cannibalizing its top-line revenue and eroding

the expected future cash flows that underpin its market valuation.

**Hypothesis 1 (Wealth Destruction and Revenue Cannibalization):** *The release of high-quality open-source software by a dominant platform fundamentally destroys the shareholder value of proprietary rivals and causes a structural decline in their sales.*

Facing severe revenue compression, the rival firm must adjust its cost structure to avoid financial distress. According to Alexy et al. (2018), the release of an open-source public good destroys the rarity and inimitability of the rival's internal resources. However, it simultaneously provides a free, high-quality technological input. The rival firm minimizes its operating costs by substituting its expensive internal development and administrative overhead ($C_{internal}$) with the new open-source infrastructure ($\Omega$). Because these survival-driven cost reductions are implemented immediately to offset the top-line shock, they mechanically inflate the firm's short-term operating margins.

**Hypothesis 2 (Defensive Free-Riding and Margin Inflation):** *To absorb the top-line shock, proprietary rivals will aggressively slash their direct costs and organizational overhead by free-riding on the open-source infrastructure, paradoxically inflating their short-term profit margins.*

## 2.4 The Surrender of Innovation and Labor Co-optation

While free-riding successfully defends short-term profitability, it profoundly alters the firm's long-term resource allocation. Let $I_R$ denote the rival's investment in exploratory proprietary research. As the focal platform continually updates the free open-source standard ($\Omega$), the marginal return on the rival's internal proprietary innovation diminishes ($\frac{\partial \Pi_R}{\partial I_R} < 0$). Consequently, the optimal strategic response is to halt internal exploratory R&D ($I_R^* \to 0$), leading to the rapid obsolescence of their existing technological portfolio.

**Hypothesis 3 (Stifling of Exploratory Innovation):** *By relying on external open-source infrastructure, proprietary rivals will systematically reduce their internal R&D investments, resulting in the decay of their exploratory innovation capabilities.*

This retreat from internal innovation imposes a strict labor reallocation constraint on the rival

firm. Let the firm's total engineering human capital be fixed such that $L_{Total} = L_{internal} + L_{external}$. To maintain product compatibility with the new industry standard and survive in the altered ecosystem, the rival is economically incentivized to shift its developers away from proprietary projects ($L_{internal}$) and direct them toward the focal platform's open-source repository ($L_{external}$) (Gawer and Henderson, 2007). By commoditizing the technological baseline, the dominant platform effectively co-opts the engineering talent of its competitors.

**Hypothesis 4 (Labor Co-optation):** *As proprietary R&D declines, incumbent rivals will actively mobilize their engineering labor force to contribute to and sustain the focal platform's open-source ecosystem.*

### 2.5 Product Market Freezing and Competitive Proximity

The ultimate consequence of this resource drain is a paralysis of the rival's external product market positioning. Divested of internal R&D funding and stripped of their engineering talent, these firms lose the capacity to dynamically update their product vocabulary or differentiate their offerings. This chilling effect effectively freezes the competitive frontier. Furthermore, standard models of product market competition suggest that the severity of these predatory dynamics—revenue cannibalization, cost-cutting, innovation decay, and labor co-optation—should not be uniformly distributed across the industry. Rather, the magnitude of the shock is strictly a function of the rival's competitive proximity to the focal platform's technological release.

**Hypothesis 5 (Targeted Freezing Effect):** *The reallocation of engineering resources will manifest as a freeze in the rival's product market fluidity. Crucially, the entirety of these predatory mechanisms will be acutely concentrated among the focal platform's most direct product market competitors.*

## 3 Data and Sample Construction

This study constructs a novel and comprehensive dataset to empirically examine the strategic impact of open-source software releases by dominant technology firms on the competitive

landscape. Given the inherently decentralized nature of open-source contributions, the data construction in this paper proceeds in three steps. Specifically, this process integrates repository-level developer activity and metadata from GitHub, textual product market network classifications from the Hoberg-Phillips database, and firm-level financial fundamentals from Compustat and CRSP.

## 3.1 GitHub Open-Source Data and Developer Engagement

To capture the open-source software releases and subsequent developer-side responses, this study constructs a highly granular dataset using the GitHub REST API and Google BigQuery.[1] The data collection and refinement process is designed to accurately isolate external competitor engagement from internal corporate development.

First, this study identifies the focal technology firms that drive the modern open-source ecosystem. These entities include top-tier technology conglomerates such as Microsoft, Alphabet, Meta, and Apple, along with major software infrastructure providers like Nvidia, Adobe, and Oracle. Although focusing on these dominant players might initially raise concerns regarding sample selection bias, the open-source software landscape is inherently characterized by extreme concentration. Foundational open-source projects and infrastructure standards are overwhelmingly initiated and maintained by these technology giants, making them the economically relevant sample for studying platform-level competitive shocks.

I systematically map these firms to their authenticated organizational accounts, such as google, tensorflow, and microsoft. To prevent the inclusion of unofficial or community-driven mirrors, I strictly match organizations using their verified corporate email domains (e.g., microsoft.com, oracle.com). Through the GitHub API, I extract comprehensive metadata for all public repositories maintained by these verified accounts. To focus on economically meaningful technological shocks, I rank these repositories by their cumulative community stars and isolate the top 100 most influential open-source projects released between January 2015 and December 2024. Table 1 provides a summary of these focal open-source releases, establishing the exogenous technological shocks that anchor the event study framework.

[1] The official documentation for the GitHub REST API is available at `https://docs.github.com/en/rest`, and the historical GitHub event logs are accessed via the GH Archive public dataset hosted on Google BigQuery, detailed at `https://www.gharchive.org/`.

Table 1 reveals the profound scale and extreme concentration of these open-source initiatives. Panel A shows that just three companies, Alphabet, Microsoft, and Meta, account for 93 of the top 100 most influential repositories[2], mobilizing millions of developer interactions. For instance, Alphabet and Microsoft each command over 1.7 million community stars and millions of direct code contributions. Panel B further highlights that even firms with fewer top-tier releases, such as Apple and Netflix, achieve massive average engagement per project, drawing tens of thousands of developers per repository. Economically, this intense level of developer traction validates the premise that these open-source releases are not mere corporate side projects but massive platform ecosystems.

*[Insert Table 1 Here]*

Figure 1 illustrates the distribution of the top 100 open-source releases across the sample period. The timeline reveals a sustained pattern of strategic open-source deployment by the focal firms, with a notable concentration in 2015 and a subsequent wave in 2019. This staggered distribution over the decade ensures sufficient variation in treatment timing, which is a fundamental prerequisite for effectively executing the stacked difference-in-differences empirical design.

*[Insert Figure 1 Here]*

Second, to observe real-time developer engagement following these releases, I leverage Google BigQuery to access the complete historical archive of GitHub event logs. I extract millions of time-series activity logs for the targeted top 100 repositories. To capture different dimensions of engagement, I categorize the raw logs into active contributions, which include direct code pushes, pull requests, and issue tracking, and passive contributions, which consist of stars and watches indicating broader community interest.

To validate the economic magnitude of focusing exclusively on these top-tier repositories, Figure 2 plots the average cumulative active and passive contributions over the first 300 days following a repository's release. Comparing the focal top 100 sample against the broader universe of other corporate open-source releases reveals a stark divergence. While typical corporate repositories accumulate modest engagement, the top 100 big tech releases rapidly absorb tens of thousands of

[2]For a complete list of the top 100 open-source repositories included in this analysis, along with their release dates and cumulative engagement metrics, see Appendix Table A1.

active code contributions and community stars within months. This exponential trajectory confirms that the selected sample represents a unique class of platform-level shocks, commanding unprecedented developer attention and validating the premise of labor reallocation.

*[Insert Figure 2 Here]*

A critical empirical challenge in measuring competitor response is linking individual developers on GitHub to their respective employing firms. To resolve this, I extract email domains and self-reported company names from developer profiles and commit logs, matching them directly to the official web domains and corporate names of public firms. These matched entities are then linked to standardized firm identifiers (GVKEY) in the CRSP and Compustat databases. To ensure the absolute integrity of the corporate affiliation, I strictly filter out generic personal email providers such as gmail.com or yahoo.com.

Although excluding generic domains inevitably introduces Type II errors by omitting corporate developers who contribute via personal accounts, it is a necessary empirical design to strictly minimize Type I errors, falsely attributing independent or hobbyist developer activity to a peer firm. By ensuring the absolute fidelity of the firm-developer linkage, the measured engagement provides a conservative lower bound of actual corporate involvement. Therefore, any statistically significant treatment effects observed under this stringent criterion represent a robust, conservative estimate that likely understates the true magnitude of the competitive shock on labor reallocation.

Finally, to ensure that the measured engagement reflects genuine external involvement rather than the focal firm's internal R&D activity, I impose a strict self-loop filter. Specifically, I exclude any GitHub activity generated by developers employed by the focal firm that released the repository. For instance, contributions by Microsoft employees to a Microsoft-owned repository are entirely removed from the aggregate engagement metrics. This rigorous filtering ensures that the resulting panel data cleanly captures the external labor mobilization of other firms shifting their human capital toward the newly released open-source commons.

## 3.2 Firm Financials, Market Valuation, and Innovation Outcomes

To evaluate the financial and operational consequences of open-source entry, I construct a comprehensive firm-year panel using accounting, market, and intellectual property data. I obtain fun-

damental annual data from the Compustat database via Wharton Research Data Services (WRDS) for the period spanning 2010 to 2024. To incorporate market valuations and stock performance, I merge the accounting data with CRSP pricing files using the CRSP Compustat Merged link table, accurately mapping Compustat GVKEY to CRSP PERMNO identifiers. Following standard empirical practices in corporate finance, I exclude financial institutions and utility firms, as their capital structures and regulatory environments differ fundamentally from those of technology and industrial firms.

Using this integrated panel, I compute a suite of key operational and financial variables designed to capture both top-line impacts and defensive corporate adjustments. To measure profitability and operational efficiency, I calculate Sales, Return on Assets, EBIT margin, and EBITDA margin. To directly observe the survival-driven restructuring mechanisms hypothesized in this study, I construct detailed cost structure variables, specifically Cost of Goods Sold and Selling General and Administrative expenses, both scaled by total sales. Additionally, to mitigate omitted variable bias and ensure strict covariate balance, I construct a standard set of time-varying firm controls, including Firm Size, the Market-to-Book ratio, and Leverage (Fama and French, 1992; Frank and Goyal, 2009; Roberts and Whited, 2013).

Furthermore, to assess the long-term strategic shift in corporate research and development, I enrich the financial panel with patent application data from the United States Patent and Trademark Office. Building on the innovation data framework established by Kogan et al. (2017), I construct dynamic measures of innovation quality and technological trajectory.

Specifically, following the established literature on corporate innovation (Manso, 2011; Balsmeier et al., 2017), I classify patents into explorative and exploitative categories based on their Cooperative Patent Classification (CPC) codes. Explorative patents are defined as those filed in technological classes that are entirely new to the firm, capturing radical innovation at the technological frontier. In contrast, exploitative patents represent incremental improvements within the firm's existing technological domains. Using these standard classifications, I compute the ratio of exploratory patents to the total patent output to capture the firm's risk-taking in R&D. Finally, I calculate the average age of the firm's overall patent portfolio and the specific age of its exploratory patents. These precise innovation metrics allow this study to empirically test whether the commoditization shock from big tech open-source releases stifles independent discovery and forces rival firms into

an aging, defensive innovation trajectory (Benner and Tushman, 2002).

## 3.3 Text-based Product Market Networks and Fluidity

To accurately map the competitive landscape and identify product market peers, this study relies on the text-based datasets developed by Hoberg and Phillips.[3] Unlike traditional static industry classifications such as SIC or NAICS, the Text-based Network Industry Classifications (TNIC) compute pairwise firm similarity scores based on the computational text analysis of product descriptions extracted from annual 10-K filings (Hoberg and Phillips, 2016). This continuous measure of product similarity provides a highly granular and dynamic mapping of competition, which is particularly critical in the fast-evolving technology sector where product boundaries frequently overlap and shift. I utilize these TNIC pairwise similarity scores to systematically identify the direct and indirect product market rivals of the focal big tech firms at the time of their open-source releases.

In addition to peer identification, I extract measures of product market fluidity and product market scope from the same database to use as structural outcome variables. Product market fluidity captures the extent to which a firm's product vocabulary is changing relative to the aggregate changes in its competitors' vocabularies, serving as a powerful proxy for product space dynamism and competitive threat (Hoberg et al., 2014). Product market scope measures the overall breadth of a firm's product offerings. I employ fluidity and scope as key dependent variables to empirically test whether the open-source entry by dominant platforms induces a chilling effect that freezes the product development cycles of incumbent peers.

By integrating the GitHub developer activity, Compustat financial metrics, USPTO patent characteristics, and the Hoberg-Phillips product market network, I construct a multidimensional firm-year panel. The TNIC similarity scores specifically lay the essential groundwork for the empirical identification strategy. The exact methodology for utilizing these text-based scores to assign treatment status, categorize competitive threat levels, and construct the optimally matched cohorts for the stacked difference-in-differences framework is detailed in the following empirical strategy section.

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

## 3.4 Treatment Assignment and Cohort Construction

To rigorously identify the causal impact of big tech open-source releases, I employ a stacked difference-in-differences framework. The critical first step in this empirical design is defining exactly which firms are exposed to the technological shock. Because open-source releases consist of underlying code and digital infrastructure, the natural candidates for exposure are firms operating within the broader technology and software sectors. However, a general technological release does not affect all software firms uniformly. The true economic impact is dictated by a firm's specific market positioning. Ultimately, the firms that face the most severe competitive shock are those providing similar proprietary products or services, as the free provision of a substitute technology by a dominant platform directly threatens to commoditize their core business. Therefore, accurately defining the treatment group requires a granular mapping of overlapping product spaces rather than relying on broad, static industry groupings.

To achieve this precise identification, I utilize the Hoberg-Phillips TNIC pairwise similarity scores to identify the product market competitors of the focal releasing firm. This text-based measure captures the exact degree of product market overlap based on the detailed product descriptions in annual 10-K filings. A primary empirical concern when using such text-based metrics is look-ahead bias, where the open-source release itself might force a peer firm to pivot its business model and subsequently alter its future product descriptions. To eliminate this endogeneity, I assign treatment status relying strictly on the TNIC network at year $t-1$, the year immediately preceding the open-source release. Any firm possessing a non-zero TNIC similarity score with the focal big tech firm in this pre-event year is designated as an exposed peer in the treatment group.

To further unpack the cannibalization mechanism and explore treatment heterogeneity, I stratify the treated firms into three distinct competitive threat tiers based on their pre-event TNIC scores. Specifically, within each focal firm's event cohort, I rank the treated peers by their pairwise similarity to the focal firm and classify them into High, Mid, and Low Threat terciles. Appendix Table IA.2 presents a representative sample of these matched peers across the major focal firms.

This classification demonstrates strong face validity for core technology markets. For instance, Nvidia's High Threat peers accurately identify direct competitors in the semiconductor and electronic equipment sectors, such as Marvell Technology and Qualcomm. Similarly, Alphabet's High

Threat group captures direct rivals in the digital advertising and online platform space, including Yelp and ReachLocal, while Apple's High Threat peers effectively capture major software developers like Adobe.

Admittedly, the strict reliance on automated text analysis occasionally introduces cross-industry noise, particularly for focal firms with highly diversified or idiosyncratic product vocabularies in their regulatory filings. As illustrated in the appendix table, Meta and Cloudflare exhibit anomalous matches with firms in the energy and heavy manufacturing sectors. However, maintaining this objective, data-driven assignment is a critical empirical design choice. It strictly binds the analysis to the ex-ante algorithmic outputs and completely removes any researcher discretion or cherry-picking.

From an econometric perspective, such cross-industry noise functions as classical measurement error in the treatment assignment, which inherently biases the estimated coefficients toward zero through attenuation bias. Consequently, documenting statistically significant cannibalization effects across these threat tiers, despite the presence of mismatched peers, strongly implies two things. First, the threat stratification successfully captures genuine varying degrees of product market overlap. Second, the baseline empirical estimates are highly robust and represent a conservative lower bound of the true competitive impact of big tech open-source releases.

Because treated firms in the TNIC network may possess fundamentally different risk profiles and growth trajectories compared to the broader Compustat universe, comparing them directly to all untreated firms introduces severe selection bias. To establish a valid counterfactual, I execute a one-to-one nearest-neighbor Propensity Score Matching at $t-1$. The control pool consists of all Compustat firms outside the focal firm's TNIC network in the pre-event year. I estimate the propensity scores using the fundamental firm characteristics defined earlier, including Firm Size, the Market-to-Book ratio, Return on Assets, and Leverage. This matching procedure forces the control group to perfectly mirror the observable pre-treatment financial conditions of the treated peers, isolating the treatment effect from pre-existing firm distress or divergent industry life cycles.

Table 2 reports the summary statistics for the final firm-year panel, partitioned into the treated and matched control groups. A comparison of Panel A and Panel B demonstrates the effectiveness of the propensity score matching procedure. Fundamental financial characteristics, including Firm Size, the Market-to-Book ratio, Return on Assets, and Leverage, are highly comparable across both

cohorts. This covariate balance confirms that the matched control sample provides a valid, structurally similar counterfactual, mitigating concerns that unobserved pre-existing financial distress or divergent growth trajectories might drive the empirical results.

Beyond financial parity, the summary statistics reveal a stark divergence in human capital mobilization. Despite operating under similar financial conditions, treated firms exhibit substantially higher levels of GitHub engagement compared to their matched controls. For instance, the treated cohort averages 25.72 total GitHub contributions and 6.26 active developers per firm-year, which is roughly double the engagement metrics observed in the control group. Economically, this descriptive divergence provides compelling preliminary evidence for the hypothesized mechanism. It confirms that firms occupying the identical product space as the focal big tech entity are significantly more likely to reallocate their engineering resources toward the newly released open-source repository, validating the text-based TNIC assignment as a precise measure of technological exposure.

*[Insert Table 2 Here]*

## 3.5 Empirical Strategy: Stacked Difference-in-Differences

To identify the causal impact of big tech open-source releases, 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 effects are heterogeneous across cohorts (Baker et al., 2022). The final dataset comprises approximately 50,000 firm-year observations, constructed by appending the individual event cohorts into a single master panel. To ensure comparability across heterogeneous firms and to mitigate the influence of extreme outliers, continuous financial variables are systematically normalized, typically scaled by total assets or total sales, and winsorized at the 1st and 99th percentiles.

The primary empirical specification is modeled as follows:

$$Y_{i,c,t} = \alpha + \beta(Treat_{i,c} \times Post_{c,t}) + \Gamma X_{i,t} + \mu_{i,c} + \lambda_{c,t} + \varepsilon_{i,c,t} \tag{3}$$

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 high-impact open-source release event, spanning an observation window from $t = -3$

to $t = +3$ years. 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 firm, representing the exposed competitor. $Post_{c,t}$ is an indicator for the period following the open-source release where $t \geq 0$.

The vector $X_{i,t}$ represents the set of time-varying fundamental controls, including Firm Size, the Market-to-Book ratio, Return on Assets, and Leverage, ensuring that the estimates are not driven by shifting financial conditions. I rigorously control for unobservable confounders by including cohort-firm fixed effects, $\mu_{i,c}$, to absorb time-invariant firm characteristics within each specific event stack. Furthermore, I include cohort-time fixed effects, $\lambda_{c,t}$, to 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.

Following the strategic logic of commoditizing the complement, I further investigate the mechanism of cannibalization by decomposing the treatment effect across varying degrees of competitive intensity. I stratify treated firms into High, Mid, and Low Threat groups based on their TNIC pairwise similarity scores to the focal releasing firm prior to the event (Casadesus-Masanell and Llanes, 2011). This granular approach allows this study to test whether the predatory impact of open-source entry is concentrated among the most direct competitors, as predicted by the theory of information-driven commoditization, while peripheral firms might instead experience pure technological spillovers.

# 4 Empirical Results

Before detailing the empirical findings, I outline the overarching hypotheses that guide this analysis. If dominant technology platforms strategically release open-source software to commoditize complementary markets, the initial shock should manifest as a negative reassessment of expected future cash flows for incumbent proprietary rivals. Faced with this sudden revenue compression, I hypothesize that these treated firms will be forced into a defensive posture, significantly cutting operational costs to artificially maintain short-term profit margins. Consequently, this survival-driven restructuring is expected to freeze their product market positioning, resulting in a measurable chilling effect on product differentiation. As proprietary development stalls, I

predict that these firms will reallocate their engineering human capital toward the newly released open-source commons to ensure compatibility or survival. Ultimately, this structural shift should lead to the stagnation of their independent innovation capabilities, characterized by an aging patent portfolio and a sharp decline in exploratory research. The following subsections systematically test these sequential predictions.

## 4.1 Market Reaction and Wealth Destruction

I begin the empirical analysis by investigating whether the stock market anticipates the predatory nature of these open-source releases. Table 3 reports the cumulative abnormal returns for competitor firms across various event windows. To ensure the robustness of the market valuation shock, I present the results using three distinct specifications, which include the raw return, the market model, and the Fama-French four-factor model.

In column (1), I first examine the pre-event window of [-5, -2]. Across all three specifications, I find that the cumulative abnormal returns are statistically indistinguishable from zero. This lack of significance is a critical econometric condition. It confirms that there is no anticipatory price movement or information leakage prior to the actual open-source release, validating the exogeneity of the event timing.

Next, I analyze the post-period windows to observe the market's reaction following the open-source entry. Looking at the raw returns in the narrow [-1, 1] event window in column (2), I find a 1.31 percentage point drop, statistically significant at the 1% level. On an annualized basis, this immediate reaction translates to a pronounced 66.81% decline, highlighting the severity of the initial market shock. While raw returns capture this immediate panic, they fail to account for broader market movements over longer horizons, which explains the loss of statistical significance in columns (4) and (5) for the raw return specification. To isolate the true abnormal performance, I rely on the market model and the Fama-French four-factor model.

Using the market model, the magnitude and persistence of the wealth destruction become remarkably clear. In column (2) for the [-1, 1] window, I find a 0.36 percentage point drop, statistically significant at the 1% level, which represents a severe annualized loss of 26.01%. Moving across the columns, this negative shock intensifies monotonically. In column (3) for the [-1, 5]

window, the drop is 0.47 percentage points, statistically significant at the 1% level. By column (5) for the [-1, 50] window, the mean cumulative abnormal return falls by 1.89 percentage points, statistically significant at the 1% level.

The Fama-French four-factor model confirms this consistent and monotonic deterioration in peer valuations. In column (2), I find a 0.15 percentage point drop, statistically significant at the 10% level, translating to an annualized decline of 11.97%. This expands to a 0.29 percentage point drop in column (3), statistically significant at the 5% level. Ultimately, in column (5), the wealth destruction deepens to a 1.63 percentage point drop, statistically significant at the 1% level.

[Insert Table 3 Here]

To visually corroborate these tabular findings, Figure 3 traces the daily cumulative average abnormal returns for the treated cohort over an extended event window using the Fama-French four-factor model. The graphical evidence provides a compelling summary of the market reaction. Prior to the open-source release at day zero, the pre-trend is remarkably flat and hovering around zero, visually reinforcing the exogeneity of the shock and the absence of prior information leakage. Immediately following the event, the abnormal returns exhibit a sharp, persistent, and monotonic downward trajectory throughout the 50-day post-period. This continuous wealth erosion visually confirms that the market progressively reprices the incumbent firms downward as the competitive implications of the open-source entry are fully absorbed.

[Insert Figure 3 Here]

Economically, a 1.63 percentage point decline in market capitalization over a 50-day trading window, which is equivalent to an annualized loss of 7.64%, represents a massive destruction of shareholder wealth for incumbent competitors. More importantly, this persistent repricing reveals a stark consensus among investors. The market clearly interprets the open-source release not as a benign technological spillover, but as a credible, structural threat to the future viability of competing firms. This decisive investor signal strongly validates the narrative that dominant technology platforms successfully deploy open-source software as a predatory weapon to commoditize complementary markets and erode the expected future cash flows of their rivals (Casadesus-Masanell and Llanes, 2011; Nagle, 2019).

## 4.2 Does the Open-Source Shock Cannibalize Rival Revenues?

The stock market's pessimistic repricing suggests a fundamental deterioration in the peers' top-line performance. Table 4 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, starting from a baseline model to a fully saturated specification.

In column (1), using a baseline difference-in-differences specification without fixed effects, I find that normalized sales decreased in the treated group during the post-period compared to the control group, statistically significant at the 1% level. In column (2), after introducing fundamental firm-level controls, the coefficient remains stable and statistically significant at the 1% level.

To rigorously absorb unobservable confounders, I introduce strict cohort-firm and cohort-time fixed effects in column (3). Under this saturated model, I find that sales decreased by 3.8 percentage points in the treated group during the post-period compared to the control group, statistically significant at the 1% level. Furthermore, in column (4), I adjust the standard errors by clustering at the firm level to account for within-firm serial correlation over time. The negative treatment effect remains robust, statistically significant at the 10% level.

Economically, a 0.038 reduction in sales scaled by total assets is highly substantial. Relative to the sample mean of 0.724, this translates to a 5.2% decline in the overall revenue generation capacity of incumbent competitors. This severe top-line compression confirms that when dominant platforms release high-quality open-source software, it acts as a direct substitute, effectively cannibalizing the market share and pricing power of proprietary alternatives.

[Insert Table 4 Here]

To further validate the causal interpretation of this revenue shock, Figure 4 plots the dynamic treatment effect on sales relative to the base year. The visual evidence strongly supports the parallel trends assumption essential for the difference-in-differences framework. In the pre-treatment years, the estimated coefficients are statistically indistinguishable from zero, indicating that there are no divergent pre-existing trends between the treated and control groups prior to the event. However, immediately following the open-source release at year zero, there is a sharp and persistent downward trajectory in peer sales. This dynamic plot confirms that the revenue compression is not

driven by prior firm distress, but is a direct and causal consequence of the focal firm's open-source entry.

[Insert Figure 4 Here]

A 5.2 percent structural decline in sales presents a substantial operational challenge to incumbent firms. This raises a natural economic question: how do these competitors absorb such a pronounced top-line contraction without experiencing immediate financial distress? The economics of open-source software provides a critical theoretical bridge. While the strategic release of open-source infrastructure by dominant platforms inherently cannibalizes the proprietary revenues of rivals (Casadesus-Masanell and Llanes, 2011), it simultaneously introduces a high-quality, free technological input into the broader market. Incumbent firms can strategically leverage these open-source public goods to eliminate research and development redundancies, increase IT productivity, and streamline software input costs (Nagle, 2019; Emery et al., 2025).

Consequently, while the primary effect of open-source entry is severe revenue compression, an analysis of the underlying cost structures reveals a highly nuanced defensive response. In the following subsection, I demonstrate how this top-line shock induces treated firms to implement strict cost-cutting measures, an operational adjustment that mechanically inflates their short-term profit margins.

## 4.3 How Do Incumbent Rivals Defend Short-Term Profitability?

The decline in overall revenue generation capacity forces incumbent firms into a severe operational restructuring. However, the economics of open-source software provides a unique mechanism for these firms to absorb the 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, high-quality 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).

To understand how peer firms leverage this free-riding opportunity to defend their bottom line, Table 5 examines the shifts in their underlying cost structures and resultant profitability metrics.

Because this strategic restructuring fundamentally begins with utilizing the technological spillover to optimize cost efficiency, I first analyze the expenditure adjustments reported in columns (1) and (2).

In column (1), I find that the cost of goods sold decreased in the treated group during the post-period compared to the control group, statistically significant at the 1% level. The estimated coefficient of -0.040 translates to a 4.0 percentage point reduction in normalized direct costs, representing a substantial 12.7% decline relative to the sample mean of 0.316. Similarly, in column (2), I find that SG&A expenses decreased in the treated group during the post-period compared to the control group, statistically significant at the 10% level. The coefficient of -0.029 implies a 2.9 percentage point drop in overhead and administrative costs, which equates to a 7.2% decline relative to the mean of 0.404. These aggressive reductions confirm that peer firms systematically slash both their variable and overhead expenses, effectively capitalizing on the technological spillovers of the open-source release to optimize their cost efficiency.

Crucially, this free-riding and subsequent cost-cutting translates into a mechanical inflation of short-term profitability. Columns (3) through (5) report these bottom-line consequences. In column (3), I find that the return on assets increased in the treated group during the post-period compared to the control group, statistically significant at the 10% level. The coefficient indicates a 4.1 percentage point improvement in overall asset profitability. This artificial margin expansion is consistent across operating metrics. In column (4), EBIT increased by 3.0 percentage points in the treated group relative to the control group, statistically significant at the 5% level. Furthermore, in column (5), EBITDA increased by 2.8 percentage points in the treated group compared to the control group, statistically significant at the 5% level.

[Insert Table 5 Here]

To further validate the causal sequence of this defensive restructuring, Figure 5 plots the dynamic treatment effects for both EBITDA and COGS. The visual evidence strongly 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 B illustrates a sharp, persistent, and downward trajectory in the cost of goods sold over the subsequent three years. Mirroring this exact timeline, Panel A displays a consistent upward trajec-

tory in EBITDA. This dynamic symmetry provides compelling evidence that the observed margin expansion is not driven by organic business growth, but is rather the direct, mechanical consequence of aggressively shrinking the firm's cost base via open-source adoption in response to the exogenous revenue shock.

[Insert Figure 5 Here]

While the baseline estimates confirm a structural shift toward leaner cost structures and artificially inflated short-term margins, this defensive restructuring raises a deeper empirical question regarding the exact mechanism of survival. If incumbent firms are indeed defending their profitability by free-riding on the focal platform's open-source infrastructure, this strategic pivot must manifest beyond mere accounting adjustments. Theoretical models of open innovation suggest that substituting internal proprietary development with external public goods fundamentally alters a firm's resource allocation (Chesbrough and Crowther, 2006; Lin and Maruping, 2022). Therefore, if this technological free-riding mechanism holds, two distinct structural consequences should follow: first, there must be direct, observable evidence of peer firms mobilizing their engineering labor to actively engage with the newly released open-source repository (Coleman et al., 2025); second, this increasing reliance on external infrastructure should lead to a measurable stagnation in their own exploratory R&D and independent innovation capabilities (Mkrtchyan et al., 2025).

## 4.4 Mechanism: Does Free-Riding Stifle Exploratory Innovation?

As established in the previous analysis, peer firms absorb the top-line shock by aggressively cutting costs and free-riding on the newly available open-source infrastructure. While this strategic pivot successfully defends short-term margins, innovation theory warns that outsourcing core technological infrastructure profoundly alters a firm's internal resource allocation. Specifically, when incumbent firms substitute proprietary development with external open-source public goods, the reliance on external knowledge often crowds out internal research and development (Cassiman and Veugelers, 2006). Consequently, this systematic divestment leads to a decay in their independent technological trajectory, forcing a structural shift away from exploratory discovery and toward defensive, exploitative innovation (Benner and Tushman, 2002). Table 6 tests this theoreti-

cal hypothesis by examining the shifts in peer firms' innovation activities following the competitive shock.

In column (1), I find that R&D expenditures scaled by total assets decreased in the treated group during the post-period compared to the control group, statistically significant at the 5% level. The estimated coefficient of -0.010 translates to a 1.0 percentage point drop in R&D intensity. Relative to the sample mean of 0.108, this represents a substantial 9.3% decline in the firm's overall research investment. This structural reduction confirms that treated firms actively scale back their proprietary technological investments once a high-quality, free alternative becomes available.

With less capital allocated to internal research, the nature of the firm's innovation portfolio fundamentally ages. In column (2), I find that the average age of the overall patent portfolio increased in the treated group during the post-period compared to the control group, statistically significant at the 1% level. The coefficient indicates a portfolio aging of 0.366 years. This aging process is particularly severe at the technological frontier. In column (3), I find that the average age of exploratory patents increased by 0.722 years in the treated group relative to the control group, statistically significant at the 1% level.

This accelerated obsolescence is driven by a wholesale abandonment of radical discovery. In column (4), I find that the share of explorative patents in the total portfolio decreased in the treated group during the post-period compared to the control group, statistically significant at the 5% level. Economically, the coefficient of -0.022 represents a massive 33.4% decline in exploratory innovation relative to the mean exploration share of 0.0659. By commoditizing the leading edge of technology, focal platforms effectively force their rivals to abandon risk-taking discovery and pivot toward defensive, exploitative improvements on legacy systems.

[Insert Table 6 Here]

To validate the causal nature of this innovation decay, Figure 6 plots the dynamic treatment effects for R&D expenditures and average patent portfolio age. The visual evidence strongly confirms the parallel trends assumption, as the pre-treatment coefficients for both variables remain flat and statistically indistinguishable from zero prior to the event. Following the open-source release, Panel A reveals a persistent downward trajectory in R&D investment over the subsequent three years. Mirroring this structural decline, Panel B displays a continuous upward climb in the aver-

age age of the patent portfolio. This dynamic evidence proves that the stifling effect on innovation is not a pre-existing trend, but a direct consequence of the focal firm's open-source entry.

[Insert Figure 6 Here]

The empirical evidence thus far paints a cohesive economic picture of platform competition. When a dominant technology firm releases a high-quality open-source substitute, peer firms suffer severe revenue compression. To survive this immediate top-line shock, they aggressively cut operating and R&D costs by free-riding on the newly available external infrastructure. However, this short-term financial defense comes at a severe long-term strategic cost. By effectively outsourcing their core technological foundations, these incumbent rivals allow their independent proprietary innovation to age and stagnate, fundamentally abandoning the technological frontier (Wen et al., 2016; Cassiman and Veugelers, 2006). This structural retreat perfectly aligns with the strategic logic of commoditizing the complement, where the focal platform successfully forces competitors into a defensive, legacy-dependent posture while securing its own dominance over future technological standards (Casadesus-Masanell and Llanes, 2011). This profound shift in corporate trajectory raises one final, overarching empirical question. If treated firms are actively slashing their internal R&D budgets and halting proprietary exploration, where exactly is their highly skilled engineering talent being redirected?

## 4.5 Mechanism: Labor Mobilization Toward the Open Commons

Labor reallocation theories in digital markets suggest that dominant platforms can effectively co-opt the proprietary talent of their rivals by setting open technological standards (Gupta et al., 2024). To maintain product compatibility and survive in the altered landscape, rivals are economically incentivized to redirect their developers to the platform's ecosystem (Lerner and Tirole, 2002). If treated firms are indeed slashing their internal R&D budgets to free-ride, this displaced engineering human capital must be reallocated toward the new open-source infrastructure.

Before examining the precise econometric estimates, Figure 7 provides compelling preliminary visual evidence of this sheer resource drain. By plotting the cumulative daily GitHub engagement for treated firms versus their matched controls, the striking divergence in active development con-

tributions immediately following the release provides undeniable evidence that rivals are heavily redirecting internal engineering resources toward the open commons.

[Insert Figure 7 Here]

To rigorously quantify this labor mobilization, Table 7 empirically tests this mechanism by examining the real-time GitHub engagement metrics of peer firms following the competitive shock. To capture the sheer volume of human capital redirected toward the open ecosystem, I first examine the overall size of the active developer base in column (4). I find that the logarithmic transformation of working developers increased in the treated group during the post-period compared to the control group, statistically significant at the 1% level. Economically, the coefficient of 0.171 implies a 29.3% increase in active development staffing dedicated to the focal open-source tasks relative to the baseline mean of 0.584. This confirms a massive structural expansion in the raw number of engineers mobilized.

Beyond the absolute expansion of the developer base, the intensity of their engagement surges dramatically. In column (1), I find that total developer engagement, measured by any work per developer, increased in the treated group during the post-period compared to the control group, statistically significant at the 1% level. The coefficient of 0.419 represents a 39.1% increase relative to the sample mean of 1.071.

Crucially, this mobilization is driven primarily by active, structural contributions rather than mere passive observation. In column (2), I find that active contributions, which include direct code pushes and pull requests, increased by 0.378 in the treated group relative to the control group, statistically significant at the 1% level. Relative to the mean of 0.756, this constitutes a staggering 50.0% surge in active development effort. Conversely, while passive contributions in column (3) also increased in the treated group compared to the control group, statistically significant at the 10% level, the economic magnitude of a 13.3% increase is substantially smaller. This divergence proves that rival firms are not merely monitoring the new technology, but are heavily investing their engineering hours into building upon it.

[Insert Table 7 Here]

To validate the causal timing of this labor reallocation, Figure 8 plots the dynamic treatment effects for both active and passive contributions per developer. The visual evidence strongly supports

the parallel trends assumption, as the pre-treatment coefficients in both panels remain statistically indistinguishable from zero. Immediately following the open-source release at year zero, Panel A demonstrates a sharp, highly significant upward trajectory in active developer contributions. This dynamic visualization confirms that the sudden mobilization of engineering talent is not an artifact of pre-existing industry trends, but a direct, causal response to the exogenous technological shock.

[Insert Figure 8 Here]

The profound labor market channel documented here highlights the ultimate predatory efficiency of big tech open-source releases. By commoditizing the technological baseline, dominant technology firms essentially co-opt the proprietary engineering talent of their rivals. Competitors are forced to continuously finance the human capital required to sustain and expand the focal platform's open ecosystem (Coleman et al., 2025). Having established the uniform cannibalization of sales, defensive cost-cutting, innovation decay, and labor mobilization, a critical question remains regarding the firm's external market positioning. If incumbent rivals are actively reallocating their engineering resources away from internal proprietary R&D, how does this internal structural shift manifest in their actual product offerings? In the next section, I utilize text-based product market measures to examine shifts in product fluidity and scope, empirically testing whether this open-source shock effectively freezes the competitive frontier.

## 4.6 Mechanism: Product Market Dynamics and the Chilling Effect

If incumbent rivals are actively reallocating their engineering resources away from internal proprietary R&D and toward external open-source ecosystems, this internal structural shift must fundamentally manifest in their actual product offerings. Theoretical models of product differentiation suggest that when firms abandon proprietary exploration to rely on standardized external infrastructure, their product evolution inevitably stagnates. To empirically capture this external market positioning, Table 8 utilizes the text-based network industry classifications developed by Hoberg and Phillips (Hoberg et al., 2014). Specifically, I examine three textual measures extracted from the product descriptions in annual 10-K filings: Product Market Fluidity, Self-Fluidity, and Product Market Scope. Product Market Fluidity gauges how rapidly a firm's product vocabulary changes relative to the aggregate shifts in its competitors' vocabularies, acting as a proxy for

product space dynamism. Self-Fluidity strictly isolates the firm's own year-over-year vocabulary changes, directly capturing internal product development and differentiation efforts. Finally, Product Market Scope measures the overall breadth of the firm's product lines. By analyzing these continuous text-based metrics, this section tests whether the open-source shock effectively freezes the competitive frontier, forcing rivals into a static product positioning.

In column (1), I find that product market fluidity decreased in the treated group during the post-period compared to the control group, statistically significant at the 1% level. Economically, the coefficient of -0.326 represents a 6.2% decline in market fluidity relative to its sample mean of 5.296. Because product market fluidity measures the extent to which a firm's product vocabulary changes relative to competitors, this reduction indicates that treated firms are falling behind the dynamic shifts in their industry.

This stagnation is entirely driven by a halt in their internal development efforts. In column (2), I find that self-fluidity decreased in the treated group during the post-period compared to the control group, statistically significant at the 1% level. The coefficient of -2.271 translates to a severe 11.1% collapse in the firm's own product differentiation efforts relative to its mean of 20.41. Interestingly, in column (3), the coefficient for product market scope is statistically indistinguishable from zero. This combination of results suggests a profound chilling effect: incumbent firms do not necessarily abandon their existing markets or shrink their product lines, but they abruptly freeze the evolution and differentiation of those existing products (Bloom et al., 2013).

[Insert Table 8 Here]

To rigorously validate the causal nature of this product market freezing, Figure 9 plots the dynamic treatment effects for both product market fluidity and self-fluidity. The visual evidence strongly supports the parallel trends assumption, as the pre-treatment coefficients in both panels remain relatively flat and statistically indistinguishable from zero prior to the event year. Immediately following the open-source release, both Panel A and Panel B exhibit a sharp, persistent downward trajectory. This dynamic visualization definitively confirms that the sudden contraction in product differentiation is not a pre-existing industry trend, but a direct consequence of the focal platform's open-source entry.

[Insert Figure 9 Here]

The empirical findings documented thus far systematically uncover the anatomy of a platform-driven commoditization shock. When dominant tech giants release foundational open-source software, they trigger a severe top-line revenue compression for incumbent rivals. To survive, these peers strategically free-ride on the new open infrastructure, aggressively slashing their direct costs and R&D expenditures, which mechanically inflates their short-term profitability. However, this defensive restructuring exacts a profound long-term strategic cost: peers suffer a severe decay in exploratory innovation, heavily reallocate their proprietary engineering talent to sustain the focal firm's open ecosystem, and ultimately freeze their product market differentiation.

Having established this comprehensive chain of cannibalization and co-optation across the average treated firm, a final layer of empirical scrutiny is required. Product market theory dictates that such predatory dynamics should not affect all competitors uniformly, but should be uniquely concentrated among the focal firm's closest product market rivals. In the next section, I conduct a rigorous robustness and heterogeneity analysis by stratifying the treated cohort by their pre-event competitive threat levels, testing whether the severity of these mechanisms strictly aligns with the proximity to the focal platform's technological frontier.

# 5 Heterogeneity Analysis by Competitive Threat Levels

The baseline results demonstrate a profound average treatment effect: focal firms' open-source releases trigger severe revenue compression, defensive cost-cutting, innovation decay, and labor co-optation among incumbent peers. However, product market theory dictates that the severity of these predatory dynamics should fundamentally depend on a competitor's proximity to the focal platform's technological frontier. In this section, I decompose the average treatment effects by stratifying the treated cohort into three distinct peer groups, Low, Mid, and High Threat, based on their pre-event product market proximity to the open-source release.

## 5.1 Are Defensive Adjustments Concentrated Among Direct Rivals?

To test whether the commoditization shock acutely targets direct competitors, Table 9 re-estimates the baseline financial impacts across the three competitive threat tiers. If open-source entry is indeed a targeted predatory weapon, the most severe operational restructuring should be

strictly concentrated in the High Threat group.

I first examine the broad market effects on revenues and direct input costs. In column (1), I observe negative coefficients for sales across the threat groups, with the Low Threat group showing a statistically significant drop of 4.8 percentage points. Similarly, in column (4), the cost of goods sold (COGS) exhibits significant reductions across all three tiers, with the High Threat group dropping by 3.7 percentage points and the Low Threat group by 4.7 percentage points. This broad dispersion suggests that big tech open-source releases act as a macro-level technological shock: they universally compress software pricing power (depressing sales) while simultaneously providing a free, generalized technological input that lowers direct manufacturing and service costs (COGS) across the entire industry. The dynamic plots in Appendix Figure IA1 (Sales) and Figure IA3 (COGS) visually corroborate that these top-line and direct-cost effects manifest broadly across the competitive landscape.

Crucially, however, the narrative sharply diverges when analyzing the firms' survival-driven organizational restructuring. While general direct costs fall across the board, aggressive reductions in overhead and administrative capacity are exclusively executed by the most vulnerable firms. In column (5), I find that SG&A expenses decreased by a massive 5.3 percentage points in the High Threat group, statistically significant at the 1% level. In stark contrast, the coefficients for the Mid and Low Threat groups are statistically indistinguishable from zero.

Because direct rivals are forced to aggressively slash these overhead and operational expenses to survive the targeted shock, their short-term profitability artificially spikes. Columns (2) and (3) confirm this highly concentrated mechanical inflation. In the High Threat group, ROA increases by 5.8 percentage points and EBITDA increases by 5.1 percentage points, both statistically significant at the 1% level. Meanwhile, the profitability metrics for the Low Threat group remain entirely flat and insignificant. The dynamic trajectory in Appendix Figure IA2 perfectly illustrates this divergence: while the EBITDA for Low Threat firms remains flat following the event, the High Threat group exhibits a steep, monotonic upward trajectory in profitability.

[Insert Table 9 Here]

These heterogeneous findings confirm a critical nuance of the platform competition mechanism. While the introduction of high-quality open-source infrastructure broadly affects industry-

wide pricing and direct input costs, the extreme, survival-driven restructuring, characterized by slashed organizational overhead and artificially inflated margins, is a localized disease. The closer a rival's product space is to the focal open-source release, the more frantically it must restructure its operations to avoid immediate financial distress.

## 5.2 Is Innovation Decay Driven by Direct Rivals?

Building on the finding that direct competitors aggressively slash organizational overhead to survive the top-line shock, it is critical to examine whether this financial squeeze spills over into their long-term technological investments. If High Threat firms are the ones facing the most severe commoditization pressure, their reliance on external open-source infrastructure and the subsequent decay in internal innovation should be the most pronounced. Table 10 tests this proposition by decomposing the shifts in innovation activities across the three competitive threat tiers.

In column (1), I find that the reduction in research and development expenditures is exclusively concentrated among the most direct product market rivals. R&D intensity decreased in the High Threat treated group during the post-period compared to the control group, statistically significant at the 5% level. The estimated coefficient of -0.016 translates to a substantial 14.8% decline in normalized R&D spending relative to the sample mean of 0.108. In contrast, the coefficients for the Low and Mid Threat groups are statistically indistinguishable from zero, indicating that peripheral competitors maintain their internal research budgets. The dynamic plot in Appendix Figure IA4 visually confirms this targeted divestment, showing a persistent and steep downward trajectory in R&D exclusively for the High Threat cohort following the open-source release.

Because these direct rivals are cutting their proprietary research budgets to free-ride on the new open infrastructure, their existing innovation pipelines age rapidly. Columns (2) and (3) illustrate this accelerating obsolescence. While the overall patent portfolio age increases for both Mid and High Threat groups, the stagnation at the technological frontier is most severe for direct competitors. In column (3), the average age of exploratory patents increased by 0.896 years in the High Threat group relative to the control group, statistically significant at the 1% level. This aging effect monotonically intensifies as the competitive threat level rises, growing from 0.461 in the Low Threat group to 0.896 in the High Threat group.

Finally, in column (4), the share of exploratory innovation exhibits negative coefficients across all tiers, reflecting a general industry-wide retreat from the technological frontier. Although statistical power is diluted when slicing the sample, the economic magnitude remains large for the High Threat group with a coefficient of -0.023. Together, these results paint a consistent picture. The stifling effect on exploratory innovation is not evenly distributed. It is the direct product market rivals, desperate to defend their short-term profit margins, that aggressively hollow out their R&D departments and allow their proprietary technologies to stagnate.

[Insert Table 10 Here]

This targeted hollowing out of internal R&D naturally leads back to the question of human capital. If direct competitors are the only ones systematically halting their proprietary research, they should also be the primary drivers of the massive labor reallocation observed in the previous sections. In the following subsection, I test whether the sudden mobilization of developers toward the GitHub open-source ecosystem is similarly concentrated among these High Threat firms.

## 5.3 Direct Rivals' Human Capital

If direct product market rivals are exclusively hollowing out their internal R&D budgets to survive the open-source shock, this displaced engineering talent must be reallocated somewhere. Building on the labor co-optation theory, it is highly probable that the exact firms abandoning their proprietary technological frontier are the ones most aggressively redirecting their developers toward the focal platform's open ecosystem. Table 11 empirically tests this dynamic by decomposing the real-time GitHub engagement metrics across the three competitive threat tiers.

The results reveal a striking concentration of labor mobilization. Starting with the absolute expansion of the workforce dedicated to the open-source project, column (4) shows that the developer base expansion is overwhelmingly driven by the High Threat group. The coefficient for the High Threat treated group is 0.302, statistically significant at the 1% level. Relative to the baseline mean of 0.584, this implies a massive 51.7% increase in active development staffing.

This dramatic expansion in headcount is matched by an explosion in actual coding effort. In column (2), I find that active contributions from the High Threat group surged by 0.761 relative to the control group, statistically significant at the 1% level. When compared to the sample mean of

0.756, this represents an astonishing 100.6% increase in direct development efforts, such as code pushes and pull requests. In stark contrast, the Mid Threat group shows absolutely no significant change in active contributions, and while the Low Threat group exhibits a marginal increase, its economic magnitude is less than half that of the High Threat cohort.

[Insert Table 11 Here]

To visually anchor these tabular findings, Figure 10 illustrates the cumulative active and passive GitHub engagement across the threat groups. The graphical divergence is profound. Immediately following the open-source release, firms in the High Threat group experience the steepest and most significant surge in active development contributions, entirely separating themselves from the rest of the cohort.

[Insert Figure 10 Here]

This targeted mobilization is further corroborated by the dynamic treatment effects plotted in Appendix Figure IA5. While the active contributions for the Low and Mid Threat groups hover around zero both before and after the event, the High Threat group exhibits a sharp, persistent spike in active coding effort strictly following the open-source release at year zero.

Collectively, this heterogeneity analysis provides definitive proof of the predatory nature of big tech open-source releases. The spillover effects are not evenly distributed technical externalities. Instead, they are highly concentrated strategic weapons. The firms whose product markets are most vulnerable to the new open-source alternative are the exact firms forced to permanently slash their internal innovation capabilities and surrender their most valuable engineering human capital to sustain the dominant platform's ecosystem.

## 5.4 Heterogeneity: The Divergent Paths of Product Market Freezing

Having established that direct rivals systematically abandon their proprietary R&D and redirect their engineering talent toward the focal open-source ecosystem, a natural final question arises regarding their actual product market survival. If the developers have essentially left the building to work on the big tech platform's infrastructure, the proprietary product lines they left behind

should inevitably stagnate. Table 12 tests this chilling effect by examining the shifts in product market fluidity and scope across the three competitive threat tiers.

The text-based measures reveal a striking contrast in how firms navigate the commoditization shock. While self-fluidity in column (2) declines across all threat groups, indicating a general industry-wide slowdown in internal product vocabulary changes, the relative market positioning tells a highly targeted story. In column (1), the drop in product market fluidity is monotonically amplified as the threat level increases. The High Threat group experiences a severe 0.429 reduction in fluidity, statistically significant at the 1% level, which is substantially larger than the 0.270 reduction observed in the Low Threat group. This indicates that while everyone slows down, the direct competitors are falling behind the dynamic shifts of the broader industry at an accelerated pace. Appendix Figure IA6 visually confirms this targeted paralysis, showing that the sharpest and most persistent downward trajectory in product market fluidity belongs exclusively to the High Threat cohort.

Perhaps the most fascinating economic insight emerges from the stark divergence in product market scope reported in column (3). For the High and Mid Threat groups, the coefficients are statistically indistinguishable from zero, confirming that these vulnerable firms are completely paralyzed—they are unable to expand into new product lines while their existing products freeze. In stark contrast, the Low Threat group exhibits a highly significant 1.155 increase in product market scope. Relative to the sample mean of 12.98, this represents a nearly 8.9% expansion in product breadth.

This juxtaposition beautifully illustrates the dual nature of big tech open-source releases. For peripheral firms (Low Threat), the open-source release acts as a benign, free technological input that allows them to cheaply diversify and expand their product scope into adjacent markets. However, for the direct product market rivals (High Threat), the exact same release functions as a devastating predatory weapon. Stripped of their pricing power and bled of their engineering talent, these direct competitors are trapped in a profound chilling effect, unable to grow their product scope and entirely frozen at the competitive frontier.

[Insert Table 12 Here]

Collectively, this heterogeneity analysis provides the ultimate validation of the paper's central

mechanism. Big tech platforms do not release high-quality open-source software merely as an act of public goodwill. By strategically commoditizing a targeted technological layer, they force their closest rivals into a desperate cycle of defensive cost-cutting, innovation decay, labor surrender, and product market paralysis, thereby securing their own uncontested dominance over the digital ecosystem.

# 6 Conclusion

The modern digital economy is increasingly defined by a profound paradox: the world's most powerful, profit-driven technology monopolies routinely invest billions of dollars to develop state-of-the-art software, only to release it to the public entirely for free. While the open-source movement is frequently romanticized as a triumph of collaborative public goodwill, this study unveils a significantly more calculated, predatory reality. By integrating the economics of two-sided network effects with the resource-based view of the firm, this paper provides comprehensive empirical evidence that dominant platforms strategically deploy open-source software as a weapon to commoditize complementary markets, deliberately dismantling the proprietary resource advantages of their incumbent rivals.

The empirical anatomy of this commoditization shock is stark and sequential. When a focal tech giant releases a foundational open-source project, the stock market immediately and accurately reprices the proprietary competitors, anticipating a structural collapse in their pricing power and future cash flows. Confronted with this severe top-line revenue cannibalization, incumbent rivals are forced into a desperate, survival-driven operational restructuring. They aggressively slash their direct manufacturing costs and organizational overhead by free-riding on the newly available open-source infrastructure. While this defensive pivot mechanically inflates their short-term profit margins, creating a temporary illusion of operational efficiency, it exacts a devastating long-term strategic toll.

By outsourcing their core technological foundations to a dominant platform, these rivals systematically hollow out their independent innovation capabilities. This study documents a profound decay in proprietary research: R&D budgets contract, overall patent portfolios rapidly age, and exploratory innovation at the technological frontier plummets. Most strikingly, the engineering

human capital that once drove these proprietary advancements is aggressively redirected. Stripped of their internal R&D mandates, competitor engineers are essentially co-opted, mobilizing in massive numbers on platforms like GitHub to maintain, debug, and expand the focal firm's open-source ecosystem. Consequently, the rival's proprietary product lines are left to stagnate, resulting in a severe chilling effect that freezes their product market differentiation and dynamism.

Crucially, the heterogeneity analysis in this study definitively proves that these dynamics are not merely the byproduct of a benign, industry-wide technological spillover. The dual nature of big tech open-source releases is highly targeted. For peripheral firms operating in distant product spaces, the open-source release functions as a free technological input, allowing them to cheaply expand their product scope. However, for the direct product market rivals whose proprietary assets are most exposed to the new open-source alternative, the exact same release acts as a devastating predatory weapon. It is strictly among these direct competitors that we observe the extreme hollowing out of R&D, the massive surrender of engineering labor, and the complete paralysis of product market fluidity.

These findings carry profound implications for both academic theory and antitrust policy. In the strategic management and economics literature, this paper bridges the gap between platform design and resource devaluation, illustrating how cross-market subsidies can be weaponized to turn a rival's valuable, rare, and inimitable (VRI) assets into obsolete legacy systems. Furthermore, it highlights a critical blind spot in modern antitrust enforcement. Traditional regulatory frameworks are inherently focused on consumer harm via monopoly pricing and output restrictions. However, in the digital platform ecosystem, the weapon of choice is not predatory pricing in the core market, but the predatory free distribution of complements.

By giving away world-class software, dominant platforms avoid traditional antitrust scrutiny while successfully starving their rivals of revenue, stifling independent exploratory innovation, and monopolizing the industry's most valuable engineering human capital. Ultimately, this study demonstrates that the true cost of free open-source software is borne by the competitive ecosystem. Through the strategic commoditization of the complement, big tech platforms systematically neutralize competitive threats and secure their uncontested dominance over the future of the digital economy.

# References


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.

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.

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.

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.

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.

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

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.

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.

Emery, L, C Lim, and S Ye, 2025, The private value of open-source innovation, *SSRN Electronic Journal* .

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

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.

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

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

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, *SSRN Electronic Journal* .

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

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.

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* .

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).

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.

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.

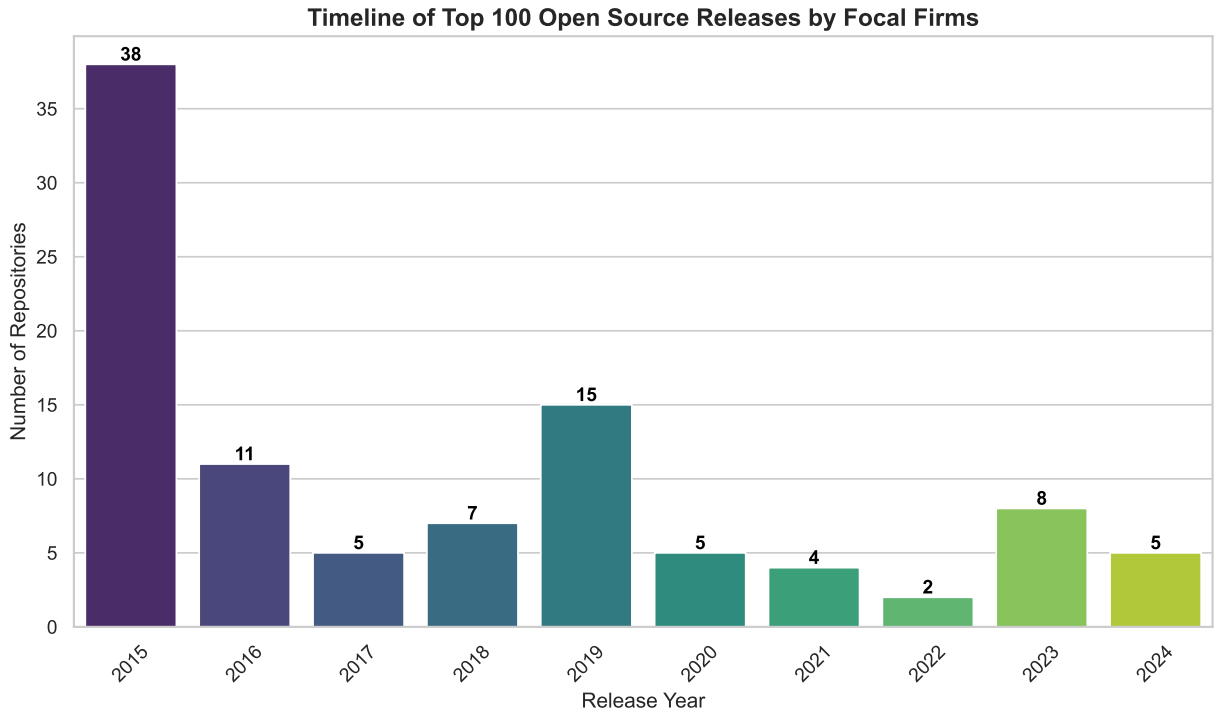



**Figure 1. Timeline of Top 100 Open Source Releases by Focal Firms**

This figure presents the annual distribution of the top 100 open-source repositories released by focal technology firms (Microsoft, Alphabet, Meta, Apple, etc.) from 2015 to 2024. The bar chart indicates the number of high-impact repositories (selected by star counts) made public each year. A significant cluster of 38 releases is observed in 2015, marking a strategic pivot toward open-source leadership by major tech giants. The subsequent steady flow of releases through 2024 demonstrates a sustained commitment to open-source as a competitive instrument, providing the temporal variation used in our stacked Difference-in-Differences (DID) estimation.

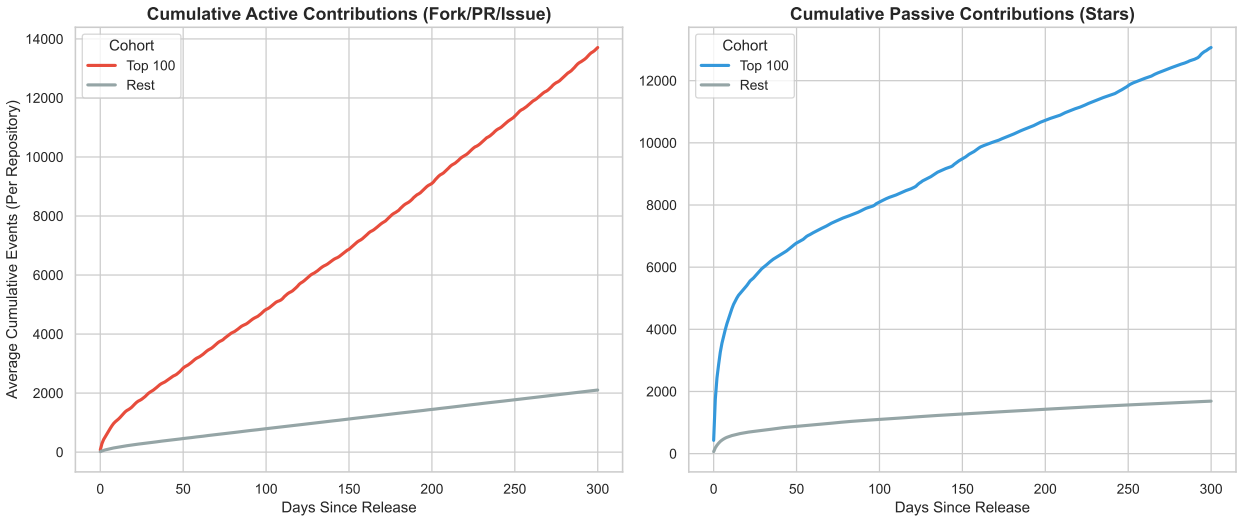



**Figure 2. Developer Engagement Trajectories: Active vs. Passive Activity**
This figure plots the daily cumulative growth of developer engagement on GitHub over a 300-day window following open-source releases. The left panel traces active engagement, defined as direct development efforts such as pushes, pull requests, and issues. The right panel traces passive engagement, which captures community interest through metrics like stars and watches.

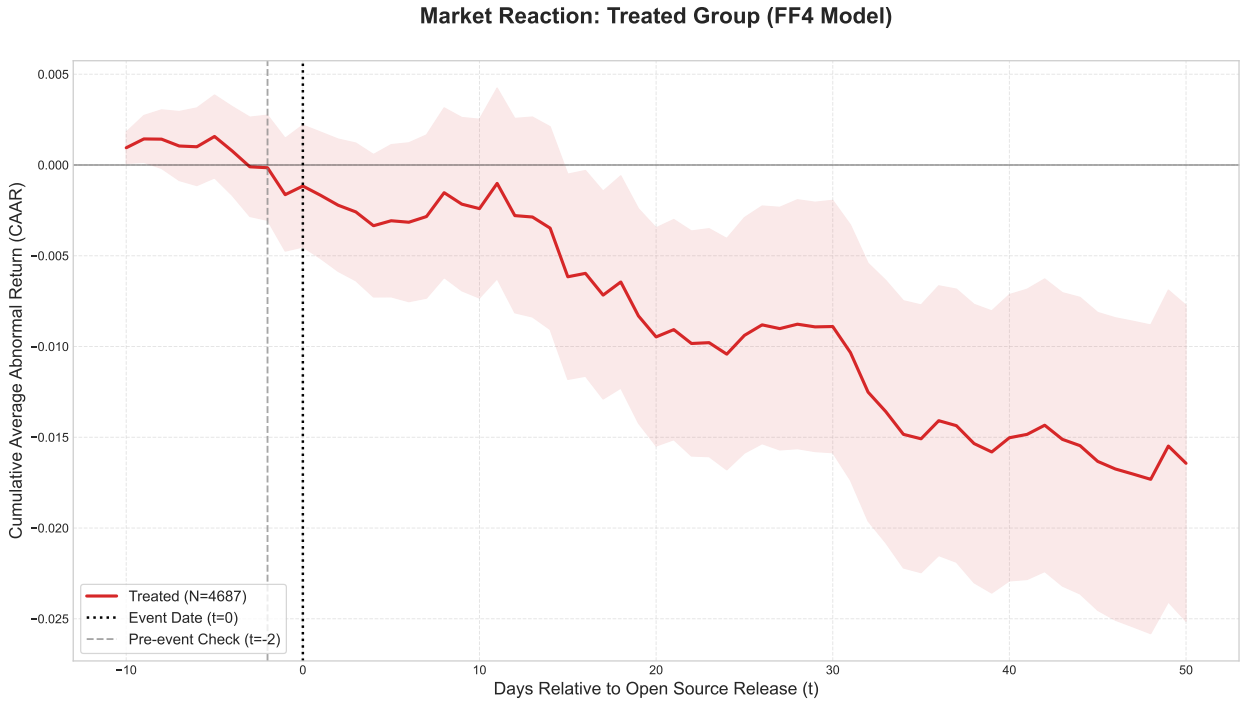



**Figure 3. Cumulative Average Abnormal Returns (FF4 Model)**
This figure displays the cumulative average abnormal returns (CAAR) for the treated group of peer firms, estimated using the Fama-French four-factor (FF4) model. The event window covers trading days [-5, 30] relative to the event day (Day 0), which is the release date of the focal firm's open-source project. The model parameters are estimated over a 150-day window ending 30 days prior to the event. The shaded area represents the 95% confidence intervals. The flat pre-trend followed by a persistent downward trajectory of the CAAR post-event confirms a significant and negative valuation shock to peer firms.

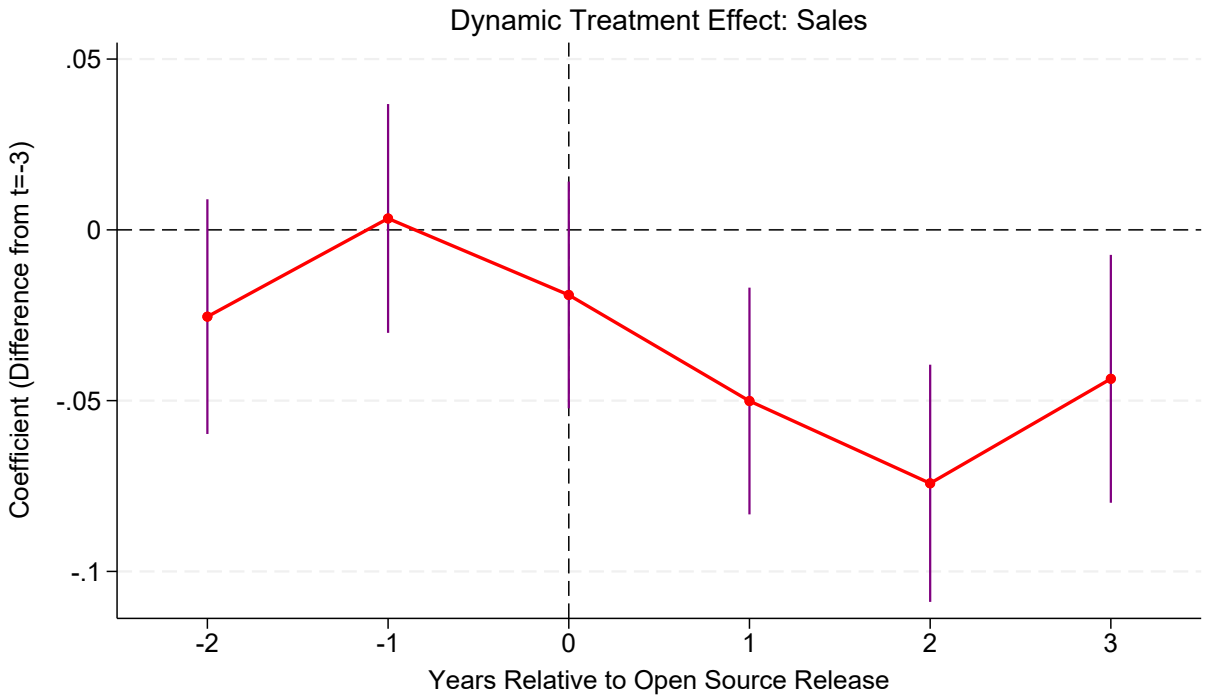



**Figure 4. Dynamic Treatment Effect on Sales**

This figure plots the dynamic treatment effect of focal firms' open-source releases on the sales of peer firms. The x-axis indicates the years relative to the open-source release (Year 0). The y-axis represents the estimated coefficients, which indicate the difference relative to the base year (t = -3). The plotted points show the annual point estimates, along with their corresponding confidence intervals.

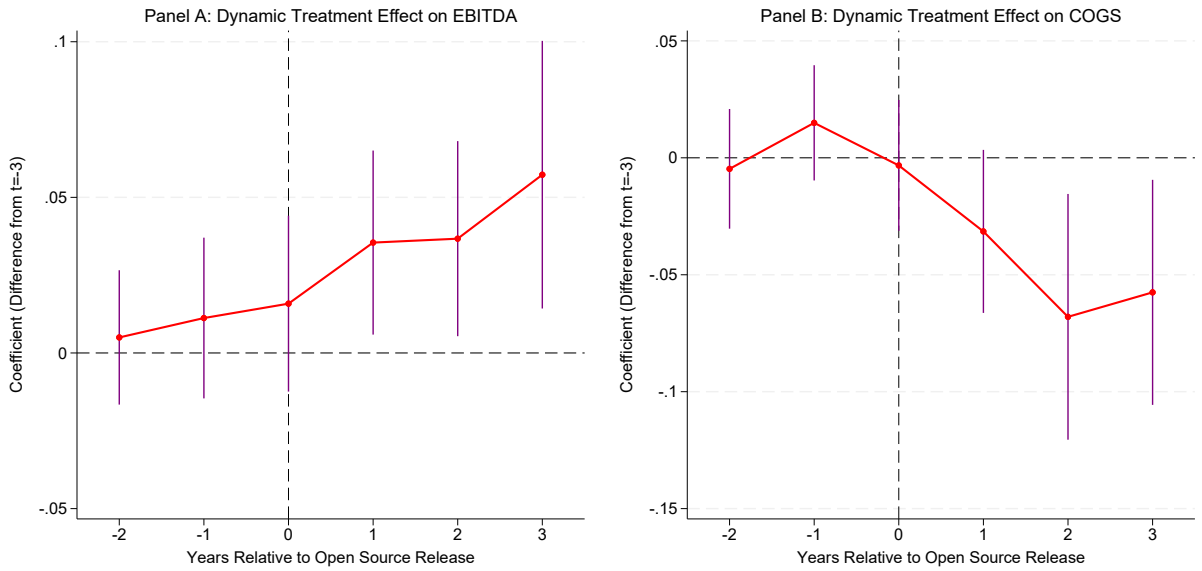



**Figure 5. Dynamic Treatment Effect on EBITDA and COGS**

This figure presents the dynamic treatment effects on the profitability and cost measures of peer firms. Panel A plots the estimated effect on Earnings Before Interest, Taxes, Depreciation, and Amortization (EBITDA), while Panel B plots the effect on Cost of Goods Sold (COGS). For both panels, the x-axis indicates the years relative to the open-source release (Year 0), and the y-axis shows the estimated coefficients relative to the base year (t = -3), accompanied by their confidence intervals.

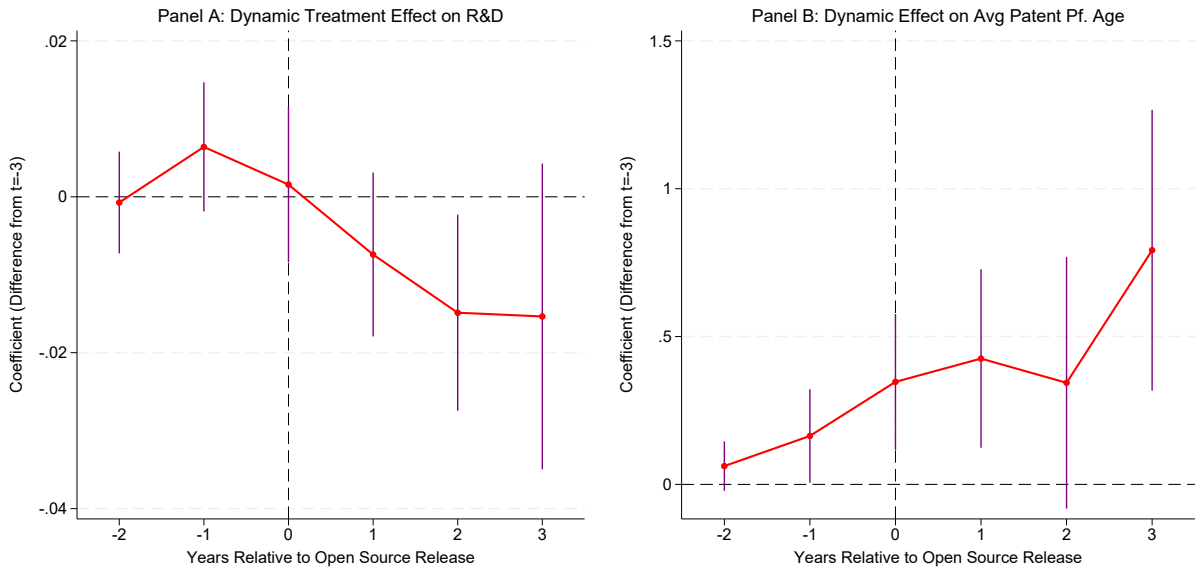



**Figure 6. Dynamic Treatment Effect on Firm Innovation**

This figure displays the dynamic treatment effects on the innovation activities of peer firms. Panel A illustrates the estimated effect on Research and Development (R&D) expenditures. Panel B shows the estimated effect on the Average Patent Portfolio Age. The x-axis represents the years relative to the focal firm's open-source release (Year 0). The y-axis displays the estimated coefficients as the difference from the base year (t = -3), with confidence intervals included for each point estimate.

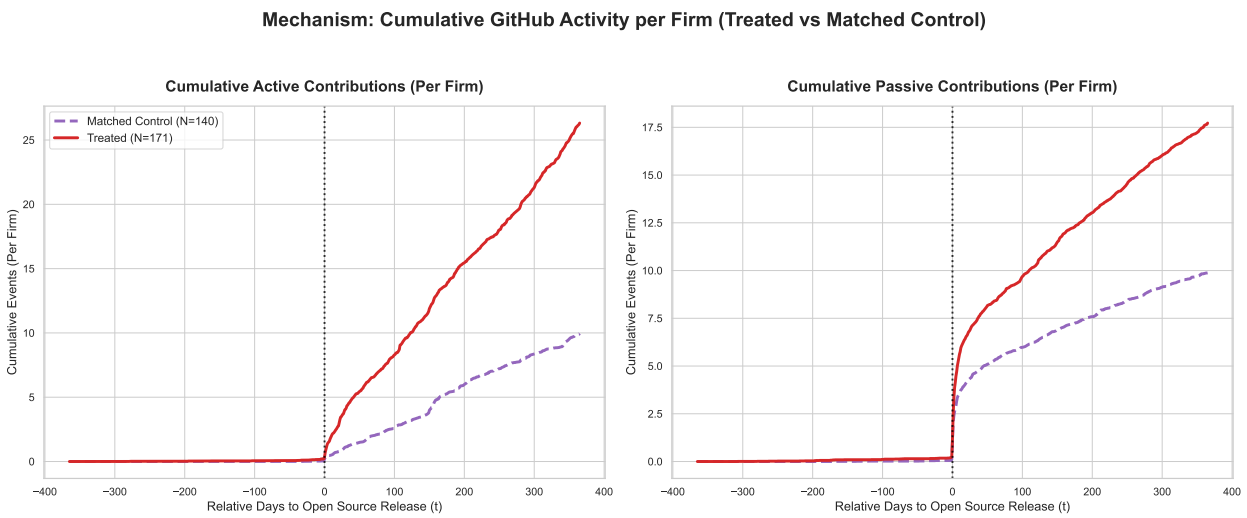



**Figure 7. Dynamics of Active vs. Passive GitHub Engagement**

This figure traces the daily cumulative growth of active versus passive engagement on GitHub for the treated firms. Active activity (red line) includes pushes, pull requests, and issues, representing direct development efforts. Passive activity (blue line) represents ”stars” and ”watches,” indicating community interest. The graph shows that while both types of engagement increase following big tech’s entry, active developer effort exhibits a steeper growth rate, suggesting a strategic shift in labor allocation toward the open-source ecosystem.

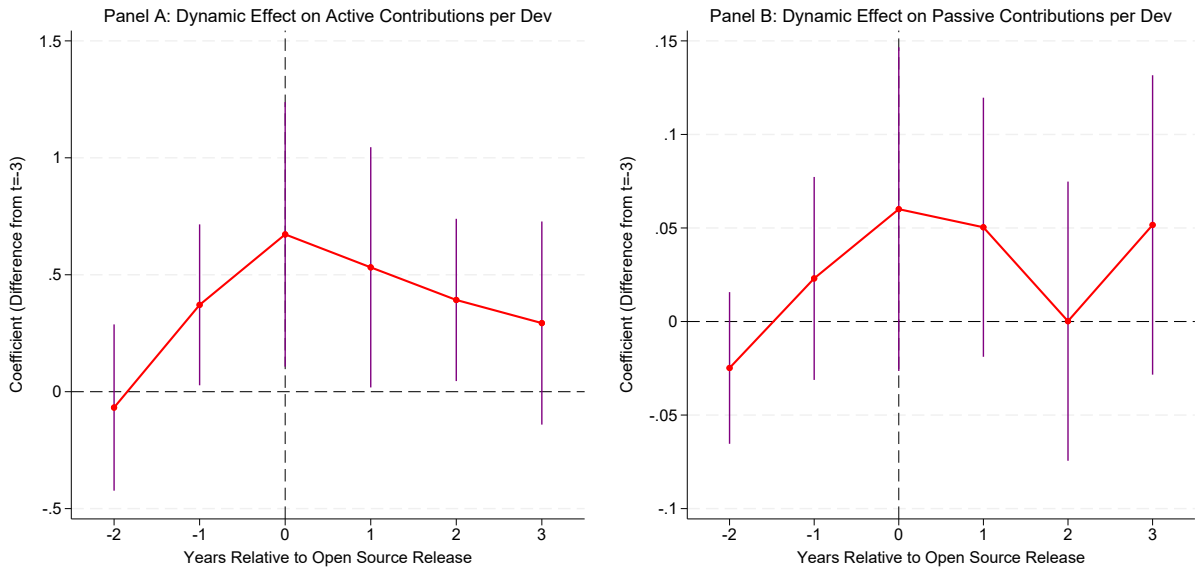



**Figure 8. Dynamic Effect on GitHub Contributions per Developer**

This figure illustrates the dynamic effects on the GitHub engagement metrics of peer firms, scaled per developer. Panel A plots the dynamic effect on Active Contributions per Developer, and Panel B plots the effect on Passive Contributions per Developer. The x-axis denotes the years relative to the open-source release (Year 0). The y-axis represents the estimated coefficients relative to the base year (t = -3), plotted with their respective confidence intervals.

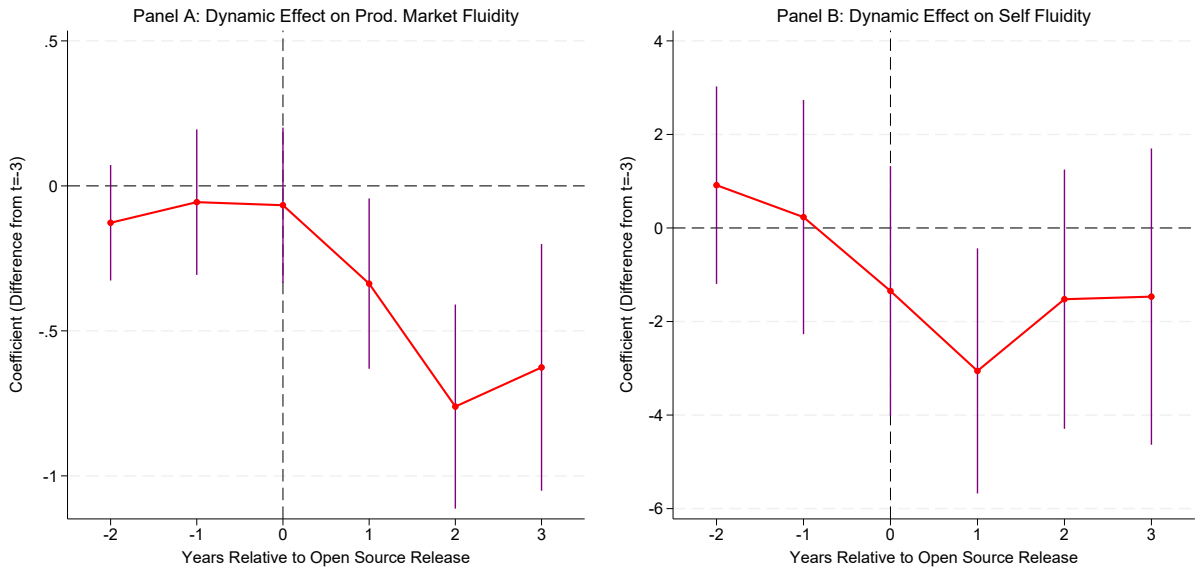



**Figure 9. Dynamic Effect on Product Market and Self Fluidity**

This figure demonstrates the dynamic effects on the product market space measures of peer firms. Panel A plots the estimated effect on Product Market Fluidity, and Panel B plots the effect on Self Fluidity. In both panels, the x-axis indicates the years relative to the open-source release (Year 0). The y-axis shows the estimated coefficients, representing the difference from the base year (t = -3), alongside their confidence intervals.

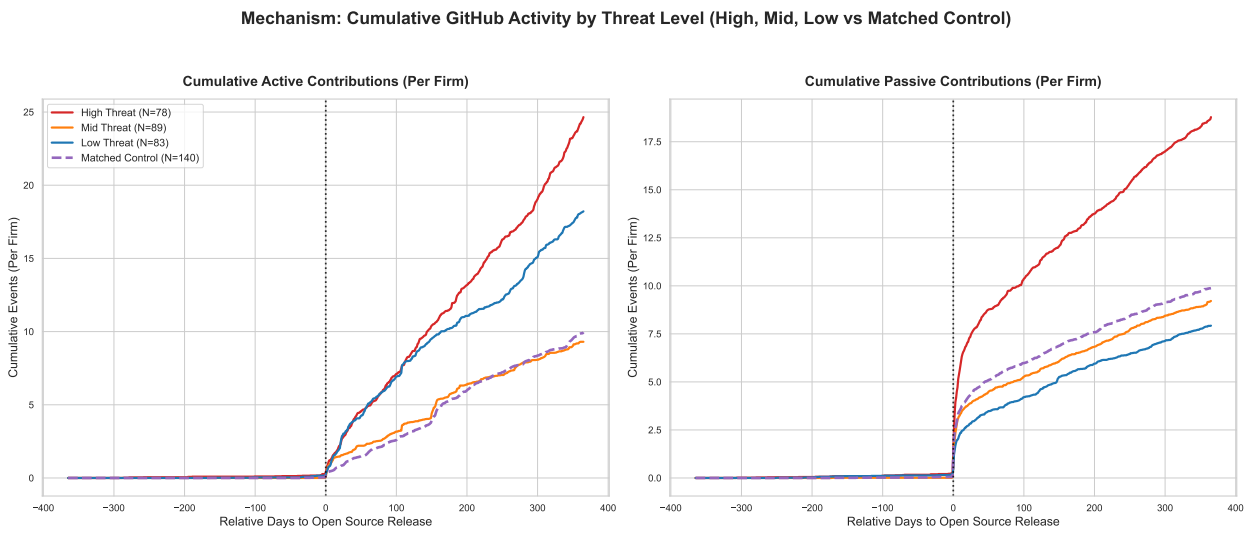



**Figure 10. GitHub Engagement Heterogeneity by Competitive Threat**

This figure presents the cumulative active (left panel) and passive (right panel) GitHub engagement across four peer groups categorized by their initial competitive threat level (High, Mid, Low, and Matched Control). The x-axis indicates the relative days surrounding the open-source release by the focal firm (Day 0). The panels illustrate that firms in the "High Threat" group experience the most significant surge in both active development contributions and passive community interest following the event. This visualizes the heterogeneity in developer engagement, indicating that the spillover effects of big tech open-source releases are concentrated in firms whose product markets are most vulnerable to the new open-source alternative.

**Table 1. Summary of Open Source Releases by Focal Company**

This table summarizes the engagement metrics across focal technology firms for their top 100 repositories from January 2015 to December 2024. Panel A presents total engagement counts, while Panel B reports average metrics per repository.

**Panel A: Total Engagement Metrics**

| Company Name | # of Repos | Total Stars | Total Forks | Total Contr. |
|---|---|---|---|---|
| Microsoft | 30 | 1,740,657 | 317,892 | 1,640,466 |
| Alphabet | 36 | 1,722,758 | 555,630 | 2,277,179 |
| Meta | 27 | 1,165,300 | 255,673 | 1,389,587 |
| Apple | 2 | 98,804 | 19,066 | 102,631 |
| Netflix | 1 | 24,237 | 5,388 | 28,271 |
| Cloudflare | 1 | 22,061 | 1,269 | 22,471 |
| Adobe | 1 | 21,495 | 7,447 | 30,681 |
| Oracle | 1 | 20,618 | 2,045 | 23,558 |
| Nvidia | 1 | 18,562 | 2,626 | 22,696 |

**Panel B: Average Engagement per Repository**

| Company Name | Avg. Stars per Repo | Avg. Forks per Repo | Avg. Contr. per Repo |
|---|---|---|---|
| Microsoft | 58,022 | 10,596 | 54,682 |
| Alphabet | 47,854 | 15,434 | 63,255 |
| Meta | 43,159 | 9,469 | 51,466 |
| Apple | 49,402 | 9,533 | 51,316 |
| Netflix | 24,237 | 5,388 | 28,271 |
| Cloudflare | 22,061 | 1,269 | 22,471 |
| Adobe | 21,495 | 7,447 | 30,681 |
| Oracle | 20,618 | 2,045 | 23,558 |
| Nvidia | 18,562 | 2,626 | 22,696 |

**Table 2. Summary Statistics: Treated vs. Matched Control Samples**

This table presents summary statistics for the main firm-year level variables used in our empirical analysis. Panel A reports the statistics for the treated sample, while Panel B reports the matched control sample. Market Capitalization and Total Assets are expressed in billions of U.S. dollars. Market-to-Book Ratio is the market value of equity divided by the book value of equity. Product Market Fluidity, Product Market Scope, and Self Fluidity measure competitive pressure and product space dynamics. GitHub variables represent counts of active/passive contributions and developer engagement. Financial variables are from Compustat. The sample period spans 2015–2024.

| | N | Mean | Std. Dev. | p10 | Median | p90 |
|---|---|---|---|---|---|---|
| **Panel A: Treated Sample** | | | | | | |
| Market Cap ($ Billions) | 27,901 | 43.83 | 190.00 | 0.10 | 2.49 | 60.67 |
| Total Assets ($ Billions) | 27,906 | 14.59 | 49.64 | 0.08 | 1.14 | 27.17 |
| Market-to-Book | 27,901 | 3.94 | 135.01 | 0.91 | 4.25 | 20.20 |
| ROA | 27,906 | -0.07 | 0.41 | -0.26 | -0.00 | 0.12 |
| Leverage | 27,906 | 0.21 | 0.21 | 0.00 | 0.15 | 0.49 |
| Product Market Fluidity | 25,598 | 5.37 | 2.06 | 2.78 | 5.33 | 7.79 |
| Product Market Scope | 23,050 | 14.30 | 8.06 | 4.00 | 13.00 | 26.00 |
| Self Fluidity | 23,910 | 19.81 | 14.66 | 6.41 | 15.72 | 38.43 |
| Active Developers | 27,911 | 6.26 | 32.15 | 0.00 | 0.00 | 11.00 |
| Total GitHub Contributions | 27,911 | 25.72 | 170.62 | 0.00 | 0.00 | 36.00 |
| Active Contributions | 27,911 | 20.08 | 154.55 | 0.00 | 0.00 | 24.00 |
| Passive Contributions | 27,911 | 5.64 | 25.93 | 0.00 | 0.00 | 11.00 |
| **Panel B: Matched Control Sample** | | | | | | |
| Market Cap ($ Billions) | 21,864 | 30.20 | 180.00 | 0.07 | 2.10 | 36.94 |
| Total Assets ($ Billions) | 21,868 | 9.17 | 36.76 | 0.06 | 1.20 | 15.06 |
| Market-to-Book | 21,859 | 5.87 | 53.91 | 0.65 | 3.55 | 14.42 |
| ROA | 21,866 | -0.07 | 0.71 | -0.28 | 0.01 | 0.14 |
| Leverage | 21,868 | 0.21 | 0.41 | 0.00 | 0.14 | 0.47 |
| Product Market Fluidity | 14,454 | 5.15 | 2.31 | 2.63 | 4.94 | 7.80 |
| Self Fluidity | 13,445 | 21.52 | 17.34 | 6.13 | 16.67 | 43.31 |
| Active Developers | 21,887 | 3.75 | 29.21 | 0.00 | 0.00 | 4.00 |
| Total GitHub Contributions | 21,887 | 12.75 | 114.34 | 0.00 | 0.00 | 8.00 |
| Active Contributions | 21,887 | 9.38 | 95.91 | 0.00 | 0.00 | 4.00 |
| Passive Contributions | 21,887 | 3.37 | 23.66 | 0.00 | 0.00 | 4.00 |

**Table 3. Cumulative Abnormal Returns (CAR) Around Open Source Entry**

This table reports the cumulative abnormal returns (CAR) for competitor firms around the announcement dates of open-source software releases by focal tech giants. We examine five different event windows: [-5, -2], [-1, 1], [-1, 5], [-1, 30], and [-1, 50]. Abnormal returns are calculated using the raw return, the market model, and the Fama-French four-factor (FF4) model. Annualized returns and t-statistics (in parentheses) are provided for each specification. The sample includes 4,687 firm-event observations. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

| | | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|---|
| **Model** | **Metric** | **[-5, -2]** | **[-1, 1]** | **[-1, 5]** | **[-1, 30]** | **[-1, 50]** |
| *Raw Return* | Mean CAR | 0.16% | -1.31%*** | -1.07%*** | 0.06% | 0.43% |
| | Annualized | 10.77% | -66.81% | -32.21% | 0.44% | 2.09% |
| | *t*-value | (1.61) | (-13.75) | (-7.82) | (0.19) | (1.17) |
| | N | 4,687 | 4,687 | 4,687 | 4,687 | 4,687 |
| *Market Model* | Mean CAR | 0.09% | -0.36%*** | -0.47%*** | -0.85%*** | -1.89%*** |
| | Annualized | 5.89% | -26.01% | -15.52% | -6.49% | -8.81% |
| | *t*-value | (0.94) | (-4.04) | (-3.39) | (-2.81) | (-4.76) |
| | N | 4,687 | 4,687 | 4,687 | 4,687 | 4,687 |
| *FF4 Model* | Mean CAR | -0.12% | -0.15%* | -0.29%** | -0.88%*** | -1.63%*** |
| | Annualized | -7.02% | -11.97% | -9.99% | -6.68% | -7.64% |
| | *t*-value | (-1.19) | (-1.80) | (-2.16) | (-2.94) | (-4.17) |
| | N | 4,687 | 4,687 | 4,687 | 4,687 | 4,687 |

**Table 4. The Impact of Big Tech Open-Source Entry on Peer Firms' Sales**

This table presents the results of a Difference-in-Differences (DiD) estimation of the effect of focal firms' open-source releases on the sales of peer firms. The dependent variable is sales divided by total assets. Columns (1) and (2) present results without fixed effects, while Columns (3) and (4) include Cohort-Firm and Cohort-Time fixed effects. Control variables include Market Capitalization, Market-to-Book (M/B) Ratio, and Leverage. Standard errors are clustered at the Cohort level in Column (3) and at the Firm level in Column (4). $t$-statistics are reported in parentheses. $^{***}$, $^{**}$, and $^{*}$ indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample period spans 2015–2024.

| | (1) Sales | (2) Sales | (3) Sales | (4) Sales |
|---|---|---|---|---|
| Treated × Post | -0.039*** | -0.037*** | -0.038*** | -0.038* |
| | (-4.24) | (-4.01) | (-4.79) | (-1.88) |
| Treated | 0.111*** | 0.112*** | – | – |
| | (16.04) | (16.12) | | |
| Post | -0.003 | 0.000 | – | – |
| | (-0.37) | (0.02) | | |
| Market Cap | | -0.133*** | 0.050*** | 0.050 |
| | | (-11.21) | (3.87) | (1.20) |
| M/B Ratio | | -0.003 | -0.003*** | -0.003 |
| | | (-1.22) | (-5.97) | (-1.57) |
| Leverage | | -0.026*** | 0.073*** | 0.073* |
| | | (-3.72) | (7.67) | (1.95) |
| Adjusted $R^2$ | 0.009 | 0.011 | 0.747 | 0.747 |
| Cohort-Firm FE | No | No | Yes | Yes |
| Cohort-Time FE | No | No | Yes | Yes |
| Clustering | No | No | Cohort | Firm |
| Mean of Dep. Var. | 0.724 | 0.724 | 0.724 | 0.724 |
| Observations | 49,772 | 49,758 | 49,758 | 49,758 |

**Table 5. Impact on Peer Firms' Profitability and Cost Structure**

This table examines the impact of focal firms' open-source releases on the profitability and cost structure of peer firms. Columns (1) and (2) report the results for cost measures: Cost of Goods Sold (COGS) and Selling, General, and Administrative (SG&A) expenses, both normalized by total sales. Columns (3) to (5) report the results for profitability measures: Return on Assets (ROA), EBIT, and EBITDA (all scaled by total assets). All specifications include control variables (Market Cap, M/B Ratio, Leverage), as well as Cohort-Firm and Cohort-Time fixed effects. Standard errors are clustered at 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.

| | Cost | | Profitability | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| | COGS | SG&A | ROA | EBIT | EBITDA |
| Treated × Post | -0.040$^{***}$ | -0.029$^{*}$ | 0.041$^{*}$ | 0.030$^{**}$ | 0.028$^{**}$ |
| | (-2.64) | (-1.75) | (1.66) | (2.49) | (2.43) |
| Adjusted $R^2$ | 0.763 | 0.742 | 0.436 | 0.714 | 0.714 |
| Controls | Yes | Yes | Yes | Yes | Yes |
| Cohort-Firm FE | Yes | Yes | Yes | Yes | Yes |
| Cohort-Time FE | Yes | Yes | Yes | Yes | Yes |
| Mean of Dep. Var. | 0.316 | 0.404 | -0.0681 | -0.0280 | 0.0132 |
| Observations | 49,758 | 48,563 | 49,758 | 49,758 | 49,717 |

**Table 6. Impact on Peer Firms' Innovation Activities**

This table estimates the impact of focal firms' open-source releases on the innovation activities of peer firms. The dependent variables are: (1) R&D expenditures scaled by total assets, (2) the average age of the patent portfolio based on filing dates, (3) the average age of exploratory patents, and (4) the share of exploratory patents in the total portfolio (Exploration %). All specifications include control variables (Market Cap, M/B Ratio, Leverage), as well as Cohort-Firm and Cohort-Time fixed effects. Standard errors are clustered at 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) R&D | (2) Patent Pf. Age | (3) Explor. Pf. Age | (4) Exploration % |
|---|---|---|---|---|
| Treated × Post | -0.010$^{**}$ | 0.366$^{***}$ | 0.722$^{***}$ | -0.022$^{**}$ |
| | (-2.07) | (2.84) | (5.05) | (-2.00) |
| Adjusted $R^2$ | 0.722 | 0.932 | 0.946 | 0.126 |
| Controls | Yes | Yes | Yes | Yes |
| Cohort-Firm FE | Yes | Yes | Yes | Yes |
| Cohort-Time FE | Yes | Yes | Yes | Yes |
| Mean of Dep. Var. | 0.108 | 4.936 | 5.903 | 0.0659 |
| Observations | 49,760 | 49,760 | 49,760 | 49,760 |

**Table 7. Mechanism: Impact on Peer Firms' GitHub Activity**

This table estimates the impact of focal firms' open-source releases on various GitHub engagement metrics of peer firms using an updated sample. Columns (1) to (3) report the results for per-developer activity (total, active, and passive contributions), while Column (4) reports the effect on the developer base size. *Active* contributions include pushes, pull requests, and issues, while *Passive* contributions include stars and watches. All specifications include control variables, as well as Cohort-Firm and Cohort-Time fixed effects. Standard errors are clustered at 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.

| | Activity per Developer | | | Developer Base |
|---|---|---|---|---|
| | (1) Any Work per Dev. | (2) Active Contrib. | (3) Passive Contrib. | (4) Ln(Working Devs + 1) |
| 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 |
| Controls | Yes | Yes | Yes | Yes |
| Cohort-Firm FE | Yes | Yes | Yes | Yes |
| Cohort-Time FE | Yes | Yes | Yes | Yes |
| Mean of Dep. Var. | 1.071 | 0.756 | 0.315 | 0.584 |
| Observations | 49,760 | 49,760 | 49,760 | 49,760 |

**Table 8. Mechanism: Impact on Peer Firms' Product Market Space**

This table estimates the impact of focal firms' open-source releases on the product market space and competitive environment of peer firms. The dependent variables are: (1) Product Market Fluidity, which measures the degree of competitive threat based on product descriptions, (2) Self Fluidity, which captures the firm's own product change and differentiation, and (3) Product Market Scope, representing the breadth of the firm's product range. All specifications include control variables (Market Cap, M/B Ratio, Leverage), as well as Cohort-Firm and Cohort-Time fixed effects. Standard errors are clustered at 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) Product Market Fluidity | (2) Self Fluidity | (3) Product Market Scope |
|---|---|---|---|
| Treated × Post | -0.326*** | -2.271*** | 0.221 |
| | (-3.13) | (-2.66) | (0.53) |
| Adjusted $R^2$ | 0.575 | 0.335 | 0.780 |
| Controls | Yes | Yes | Yes |
| Cohort-Firm FE | Yes | Yes | Yes |
| Cohort-Time FE | Yes | Yes | Yes |
| Mean of Dep. Var. | 5.296 | 20.41 | 12.98 |
| Observations | 40,015 | 37,168 | 36,036 |

**Table 9. Impact on Sales, Profits, and Costs by Competitive Threat Level**

This table examines the heterogeneous impact of focal firms' open-source releases on peer firms, categorized by their initial competitive threat level (Low, Mid, and High). Column (1) reports the impact on Sales, Columns (2) and (3) on Profitability (ROA and EBITDA), and Columns (4) and (5) on Costs (COGS and SG&A). Competitive threat is measured based on the product market proximity to the focal firm's open-source project. All specifications include control variables, as well as Cohort-Firm and Cohort-Time fixed effects. Standard errors are clustered at 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.

| | Sales | Profitability | | Cost | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| | Sales | ROA | EBITDA | COGS | SG&A |
| Treated (Low) × Post | -0.048$^{**}$ | 0.043 | 0.016 | -0.047$^{***}$ | -0.020 |
| | (-2.13) | (1.01) | (1.02) | (-2.72) | (-1.08) |
| Treated (Mid) × Post | -0.032 | 0.035 | 0.025$^{*}$ | -0.044$^{**}$ | -0.019 |
| | (-1.14) | (1.36) | (1.70) | (-2.17) | (-0.81) |
| Treated (High) × Post | -0.036 | 0.058$^{***}$ | 0.051$^{***}$ | -0.037$^{**}$ | -0.053$^{***}$ |
| | (-1.57) | (2.68) | (3.45) | (-2.13) | (-2.71) |
| Adjusted $R^2$ | 0.747 | 0.437 | 0.714 | 0.763 | 0.742 |
| Controls | Yes | Yes | Yes | Yes | Yes |
| Cohort-Firm FE | Yes | Yes | Yes | Yes | Yes |
| Cohort-Time FE | Yes | Yes | Yes | Yes | Yes |
| Mean of Dep. Var. | 0.724 | -0.0681 | 0.0132 | 0.316 | 0.404 |
| Observations | 49,758 | 49,758 | 49,717 | 49,758 | 48,563 |

**Table 10. Impact on Innovation by Competitive Threat Level**

This table examines the heterogeneous impact of focal firms' open-source releases on the innovation activities of peer firms, categorized by their initial competitive threat level (Low, Mid, and High). The dependent variables are: (1) R&D expenditures, (2) the average age of the patent portfolio, (3) the average age of exploratory patents, and (4) the share of exploratory patents (Exploration %). All specifications include control variables, as well as Cohort-Firm and Cohort-Time fixed effects. Standard errors are clustered at 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) R&D | (2) Patent Pf. Age | (3) Explor. Pf. Age | (4) Exploration % |
|---|---|---|---|---|
| Treated (Low) × Post | -0.007 | 0.154 | 0.461** | -0.021* |
| | (-1.35) | (0.95) | (2.40) | (-1.69) |
| Treated (Mid) × Post | -0.010 | 0.463*** | 0.652*** | -0.017 |
| | (-1.34) | (3.49) | (3.74) | (-1.08) |
| Treated (High) × Post | -0.016** | 0.446*** | 0.896*** | -0.023 |
| | (-2.50) | (3.17) | (5.89) | (-1.29) |
| Adjusted $R^2$ | 0.722 | 0.932 | 0.946 | 0.126 |
| Controls | Yes | Yes | Yes | Yes |
| Cohort-Firm FE | Yes | Yes | Yes | Yes |
| Cohort-Time FE | Yes | Yes | Yes | Yes |
| Mean of Dep. Var. | 0.108 | 4.936 | 5.903 | 0.0659 |
| Observations | 49,760 | 49,760 | 49,760 | 49,760 |

**Table 11. Impact on GitHub Engagement by Competitive Threat Level**

This table examines the heterogeneous impact of focal firms' open-source releases on the GitHub engagement of peer firms, categorized by their initial competitive threat level (Low, Mid, and High). Columns (1) to (3) report the impact on total (Any), active, and passive contributions per developer. Column (4) reports the impact on the developer base, measured by the natural logarithm of active developers plus one. All specifications include control variables, as well as Cohort-Firm and Cohort-Time fixed effects. Standard errors are clustered at 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.

| | Contributions per Developer | | | Developer Base |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | Any | Active | Passive | Ln(Devs + 1) |
| Treated (Low) × Post | 0.358$^{*}$ | 0.338$^{*}$ | 0.020 | 0.089 |
| | (1.70) | (1.66) | (0.62) | (1.30) |
| Treated (Mid) × Post | 0.017 | -0.007 | 0.023 | 0.138$^{*}$ |
| | (0.09) | (-0.04) | (0.73) | (1.78) |
| Treated (High) × Post | 0.848$^{***}$ | 0.761$^{***}$ | 0.086$^{**}$ | 0.302$^{***}$ |
| | (2.80) | (2.60) | (2.09) | (3.83) |
| Adjusted $R^2$ | 0.117 | 0.076 | 0.340 | 0.730 |
| Controls | Yes | Yes | Yes | Yes |
| Cohort-Firm FE | Yes | Yes | Yes | Yes |
| Cohort-Time FE | Yes | Yes | Yes | Yes |
| Mean of Dep. Var. | 1.071 | 0.756 | 0.315 | 0.584 |
| Observations | 49,760 | 49,760 | 49,760 | 49,760 |

**Table 12. Impact on Product Market Space by Competitive Threat Level**

This table examines the heterogeneous impact of focal firms' open-source releases on the product market environment of peer firms, stratified by their initial competitive threat level (Low, Mid, and High). The dependent variables are: (1) Product Market Fluidity, (2) Self Fluidity, and (3) Product Market Scope. Competitive threat is determined based on the product market proximity to the focal firm's open-source project at the time of release. All specifications include control variables, as well as Cohort-Firm and Cohort-Time fixed effects. Standard errors are clustered at 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) Product Market Fluidity | (2) Self Fluidity | (3) Product Market Scope |
|---|---|---|---|
| Treated (Low) × Post | -0.270$^{**}$ | -2.230$^{**}$ | 1.155$^{***}$ |
| | (-2.18) | (-2.13) | (2.63) |
| Treated (Mid) × Post | -0.312$^{**}$ | -1.896$^{*}$ | 0.107 |
| | (-2.51) | (-1.82) | (0.19) |
| Treated (High) × Post | -0.429$^{***}$ | -1.957$^{*}$ | -0.772 |
| | (-3.13) | (-1.65) | (-1.31) |
| Adjusted $R^2$ | 0.576 | 0.334 | 0.782 |
| Controls | Yes | Yes | Yes |
| Cohort-Firm FE | Yes | Yes | Yes |
| Cohort-Time FE | Yes | Yes | Yes |
| Mean of Dep. Var. | 5.296 | 20.41 | 12.98 |
| Observations | 40,015 | 37,168 | 36,036 |

# Internet Appendix for:
# "Commoditize Your Complement: The Cannibalization of Peer Firms by Open-Source Shocks"

# A Appendix Tables

**Table IA.1. Chronological List of Top 100 Open Source Releases**

This table provides a chronological record of the top 100 open-source repositories released by focal technology firms between 2015 and 2024. Total Stars and Total Forks reflect the cumulative engagement metrics as of December 2024. Repositories are selected based on their total star counts within each firm's official GitHub organization.

| Release Date | Company | Repository | Total Stars | Total Forks |
|---|---|---|---|---|
| 2015-01-02 | Microsoft | Microsoft/TypeScript | 98,724 | 17,110 |
| 2015-01-03 | Adobe | adobe/brackets | 21,495 | 7,447 |
| 2015-01-06 | Alphabet | google/protobuf | 29,385 | 9,284 |
| 2015-01-07 | Alphabet | google/iosched | 20,384 | 7,568 |
| 2015-01-07 | Meta | facebook/immutable-js | 24,331 | 1,867 |
| 2015-01-08 | Alphabet | angular/angular | 117,410 | 37,581 |
| 2015-01-09 | Alphabet | angular/angular.js | 43,059 | 29,601 |
| 2015-01-09 | Netflix | Netflix/Hystrix | 24,237 | 5,388 |
| 2015-01-15 | Alphabet | golang/go | 140,768 | 24,778 |
| 2015-01-16 | Meta | facebook/flow | 22,609 | 2,557 |
| 2015-01-21 | Meta | facebook/folly | 27,143 | 6,346 |
| 2015-01-21 | Meta | facebook/jest | 43,326 | 8,433 |
| 2015-02-02 | Alphabet | google/ExoPlayer | 23,176 | 7,472 |
| 2015-02-04 | Alphabet | google/dagger | 18,953 | 2,762 |
| 2015-02-09 | Meta | facebook/rocksdb | 28,241 | 6,936 |
| 2015-02-11 | Alphabet | google/fonts | 20,706 | 3,894 |
| 2015-02-12 | Alphabet | google/guava | 56,084 | 15,081 |
| 2015-02-22 | Meta | facebook/react | 266,223 | 70,077 |
| 2015-03-26 | Meta | facebook/react-native | 142,423 | 34,803 |
| 2015-03-26 | Meta | facebook/fresco | 20,076 | 5,363 |

*Continued on next page*

*Table IA1 Continued from previous page*

| Release Date | Company | Repository | Total Stars | Total Forks |
|---|---|---|---|---|
| 2015-04-01 | Alphabet | google/flatbuffers | 23,342 | 4,184 |
| 2015-04-13 | Alphabet | google/gson | 26,959 | 6,327 |
| 2015-04-25 | Microsoft | Azure/azure-quickstart-templates | 15,685 | 19,794 |
| 2015-05-13 | Alphabet | google/material-design-icons | 46,021 | 12,892 |
| 2015-05-20 | Alphabet | google/styleguide | 41,311 | 17,488 |
| 2015-06-04 | Alphabet | angular/angular-cli | 31,796 | 16,364 |
| 2015-06-11 | Meta | facebook/infer | 16,768 | 2,755 |
| 2015-07-06 | Alphabet | google/material-design-lite | 39,243 | 8,411 |
| 2015-08-10 | Alphabet | kubernetes/kubernetes | 113,262 | 47,065 |
| 2015-08-11 | Meta | facebook/relay | 20,745 | 2,499 |
| 2015-08-17 | Meta | facebook/flux | 16,723 | 5,361 |
| 2015-08-21 | Alphabet | google/googletest | 38,993 | 13,986 |
| 2015-11-04 | Nvidia | NVIDIA/nvidia-docker | 18,562 | 2,626 |
| 2015-11-09 | Alphabet | tensorflow/tensorflow | 222,055 | 113,229 |
| 2015-11-09 | Alphabet | flutter/flutter | 193,431 | 39,867 |
| 2015-11-18 | Microsoft | Microsoft/vscode | 164,552 | 27,778 |
| 2015-12-03 | Apple | apple/swift-evolution | 16,739 | 3,472 |
| 2015-12-03 | Apple | apple/swift | 82,065 | 15,594 |
| 2016-01-04 | Alphabet | angular/material2 | 19,525 | 6,152 |
| 2016-01-25 | Microsoft | Microsoft/CNTK | 17,789 | 5,202 |
| 2016-02-05 | Alphabet | tensorflow/models | 88,698 | 57,029 |
| 2016-02-22 | Meta | facebook/draft-js | 24,194 | 3,527 |
| 2016-04-15 | Alphabet | kubernetes/minikube | 31,350 | 5,918 |
| 2016-05-06 | Alphabet | google/flexbox-layout | 19,971 | 2,202 |
| 2016-06-27 | Alphabet | kubernetes/kops | 16,971 | 5,745 |
| 2016-07-22 | Meta | facebookincubator/create-react-app | 44,224 | 9,725 |
| 2016-08-04 | Meta | facebookresearch/fastText | 28,161 | 5,631 |
| 2016-08-11 | Alphabet | google/leveldb | 39,293 | 9,746 |
| 2016-09-01 | Meta | facebook/zstd | 23,885 | 2,514 |

*Continued on next page*

*Table IA1 Continued from previous page*

| Release Date | Company | Repository | Total Stars | Total Forks |
|---|---|---|---|---|
| 2017-02-21 | Alphabet | google/python-fire | 29,144 | 1,930 |
| 2017-02-28 | Meta | facebookresearch/faiss | 32,476 | 4,116 |
| 2017-04-18 | Alphabet | flutter/plugins | 19,264 | 12,204 |
| 2017-05-09 | Meta | facebookresearch/fairseq | 18,139 | 3,267 |
| 2017-10-06 | Alphabet | kubernetes/ingress-nginx | 17,430 | 9,112 |
| 2018-01-22 | Meta | facebookresearch/Detectron | 29,146 | 6,544 |
| 2018-01-22 | Meta | facebook/create-react-app | 72,792 | 30,901 |
| 2018-02-01 | Oracle | oracle/graal | 20,618 | 2,045 |
| 2018-03-31 | Alphabet | tensorflow/tfjs | 18,519 | 2,379 |
| 2018-06-08 | Alphabet | flutter/samples | 18,703 | 9,245 |
| 2018-08-03 | Alphabet | google/filament | 18,435 | 2,229 |
| 2018-11-18 | Alphabet | google/jax | 30,747 | 3,311 |
| 2019-03-06 | Microsoft | Microsoft/calculator | 33,678 | 5,250 |
| 2019-05-06 | Microsoft | microsoft/vscode | 219,500 | 60,918 |
| 2019-05-06 | Microsoft | microsoft/PowerToys | 122,816 | 8,839 |
| 2019-05-07 | Microsoft | microsoft/TypeScript | 122,588 | 19,046 |
| 2019-05-07 | Microsoft | microsoft/Terminal | 68,718 | 6,328 |
| 2019-05-14 | Microsoft | microsoft/calculator | 34,478 | 9,648 |
| 2019-05-14 | Microsoft | microsoft/vcpkg | 20,043 | 7,528 |
| 2019-05-14 | Microsoft | microsoft/terminal | 145,188 | 17,004 |
| 2019-06-16 | Alphabet | google/mediapipe | 26,152 | 5,477 |
| 2019-06-23 | Meta | facebook/docusaurus | 47,124 | 10,857 |
| 2019-06-24 | Microsoft | microsoft/monaco-editor | 27,804 | 3,050 |
| 2019-09-04 | Alphabet | google/eng-practices | 20,548 | 2,080 |
| 2019-09-18 | Microsoft | microsoft/cascadia-code | 28,595 | 1,120 |
| 2019-09-30 | Microsoft | microsoft/unilm | 20,440 | 2,730 |
| 2019-10-10 | Meta | facebookresearch/detectron2 | 32,410 | 8,868 |
| 2020-01-22 | Microsoft | microsoft/playwright | 69,346 | 4,500 |
| 2020-02-07 | Microsoft | microsoft/DeepSpeed | 35,700 | 4,386 |

*Continued on next page*

*Table IA1 Continued from previous page*

| Release Date | Company | Repository | Total Stars | Total Forks |
|---|---|---|---|---|
| 2020-05-14 | Meta | facebookexperimental/Recoil | 20,714 | 1,486 |
| 2020-05-19 | Microsoft | microsoft/winget-cli | 25,605 | 1,856 |
| 2020-11-10 | Microsoft | microsoft/Web-Dev-For-Beginners | 91,082 | 14,590 |
| 2021-03-03 | Microsoft | microsoft/ML-For-Beginners | 73,726 | 15,665 |
| 2021-03-03 | Microsoft | microsoft/Data-Science-For-Beginners | 28,586 | 6,107 |
| 2021-03-03 | Microsoft | microsoft/AI-For-Beginners | 34,361 | 6,182 |
| 2021-05-05 | Alphabet | google/zx | 34,004 | 1,202 |
| 2022-04-13 | Meta | facebook/lexical | 19,937 | 1,961 |
| 2022-12-21 | Alphabet | google/comprehensive-rust | 27,666 | 1,835 |
| 2023-02-24 | Meta | facebookresearch/llama | 49,019 | 8,188 |
| 2023-03-01 | Microsoft | microsoft/semantic-kernel | 22,163 | 3,501 |
| 2023-03-09 | Microsoft | microsoft/visual-chatgpt | 29,479 | 2,779 |
| 2023-03-31 | Microsoft | microsoft/JARVIS | 23,005 | 2,010 |
| 2023-04-05 | Meta | facebookresearch/segment-anything | 47,319 | 5,795 |
| 2023-06-09 | Meta | facebookresearch/audiocraft | 20,487 | 2,188 |
| 2023-08-18 | Microsoft | microsoft/autogen | 35,473 | 5,401 |
| 2023-10-26 | Microsoft | microsoft/generative-ai-for-beginners | 65,001 | 33,310 |
| 2024-02-27 | Cloudflare | cloudflare/pingora | 22,061 | 1,269 |
| 2024-04-13 | Microsoft | microsoft/MS-DOS | 18,194 | 3,123 |
| 2024-04-18 | Meta | meta-llama/llama3 | 26,665 | 3,108 |
| 2024-07-01 | Microsoft | microsoft/graphrag | 20,517 | 2,045 |
| 2024-12-06 | Microsoft | microsoft/markitdown | 27,821 | 1,092 |

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

This table presents representative examples of peer firms matched to the major tech giants (focal firms) based on the Hoberg-Phillips TNIC product market fluidity data. For each focal firm, this table displays two peers from the High, Mid, and Low competitive threat groups.

| Threat Level | Peer Firm Name | Broad Industry | TNIC Score |
|---|---|---|---|
| **Focal Firm: Microsoft** | | | |
| High | Polycom Inc | Electronic Equipment | 0.1384 |
| | Broadsoft Inc | Software & IT Services | 0.1167 |
| Mid | Yodlee Inc | Software & IT Services | 0.0311 |
| | Ubiquiti Inc | Electronic Equipment | 0.0308 |
| Low | Realpage Inc | Software & IT Services | 0.0002 |
| | Nuance Communications Inc | Software & IT Services | 0.0002 |
| **Focal Firm: Alphabet (Google)** | | | |
| High | Reachlocal Inc | Software & IT Services | 0.0524 |
| | Yelp Inc | Software & IT Services | 0.0302 |
| Mid | Amazon.Com Inc | Online Retail & Web | 0.0161 |
| | Logmein Inc | Software & IT Services | 0.0153 |
| Low | Chegg Inc | Software & IT Services | 0.0035 |
| | Zulily Inc | Online Retail & Web | 0.0009 |
| **Focal Firm: Meta (Facebook)** | | | |
| High | Valero Energy Corp | Other Services | 0.1721 |
| | Spectra Energy Corp | Other Services | 0.1428 |
| Mid | Regency Energy Partners Lp | Other Services | 0.0358 |
| | Northern Tier Energy Lp | Other Services | 0.0349 |
| Low | Freeport-Mcmoran Inc | Other Services | 0.0015 |
| | Seacor Holdings Inc | Other Services | 0.0015 |

**Table IA.2. Face Validity of TNIC Peer Matching for Major Focal Firms (Continued)**

| Threat Level | Peer Firm Name | Broad Industry | TNIC Score |
|---|---|---|---|
| **Focal Firm: Apple** | | | |
| High | Progress Software Corp | Software & IT Services | 0.0443 |
| | Adobe Inc | Software & IT Services | 0.0415 |
| Mid | Astea International Inc | Software & IT Services | 0.0111 |
| | Zagg Inc | Other Services | 0.0107 |
| Low | Autoweb Inc | Software & IT Services | 0.0007 |
| | Tivo Corp | Software & IT Services | 0.0002 |
| **Focal Firm: Nvidia** | | | |
| High | Marvell Technology Inc | Electronic Equipment | 0.1154 |
| | Qualcomm Inc | Electronic Equipment | 0.0818 |
| Mid | Vitesse Semiconductor Corp | Electronic Equipment | 0.0177 |
| | Omnivision Technologies Inc | Electronic Equipment | 0.0177 |
| Low | F5 Inc | Software & IT Services | 0.0005 |
| | Infinera Corp | Electronic Equipment | 0.0002 |
| **Focal Firm: Cloudflare** | | | |
| High | Trinity Industries Inc | Manufacturing | 0.1023 |
| | Martin Marietta Materials | Other Services | 0.0294 |
| Mid | Norfolk Southern Corp | Other Services | 0.0190 |
| | Gatx Corp | Other Services | 0.0090 |
| Low | Canadian Pac Kansas City Ltd | Other Services | 0.0068 |
| | Phillips 66 | Other Services | 0.0021 |

# B Appendix Figures

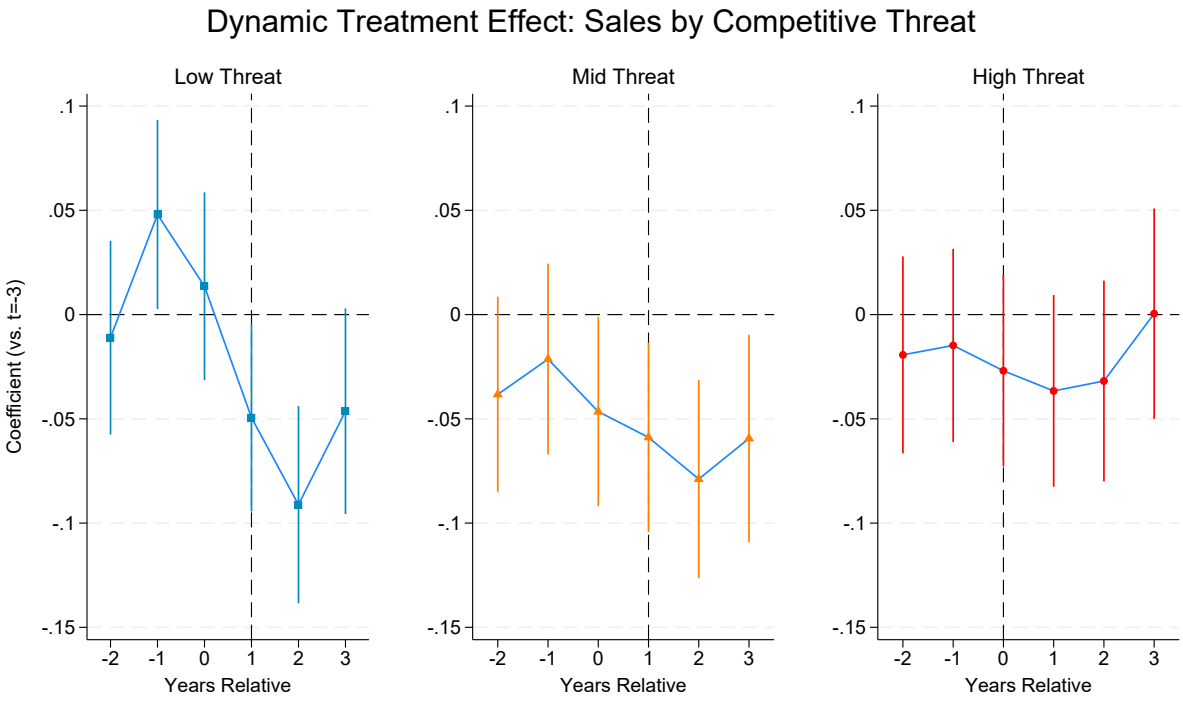



**Figure IA1. Dynamic Treatment Effect on Sales by Competitive Threat**
This figure plots the dynamic treatment effects of focal firms' open-source releases on the sales of peer firms, decomposed by their initial competitive threat levels (Low, Mid, and High Threat). The x-axis indicates the years relative to the open-source release (Year 0). The y-axis represents the estimated coefficients, indicating the difference relative to the base year ($t = -3$), plotted alongside their 95% confidence intervals.

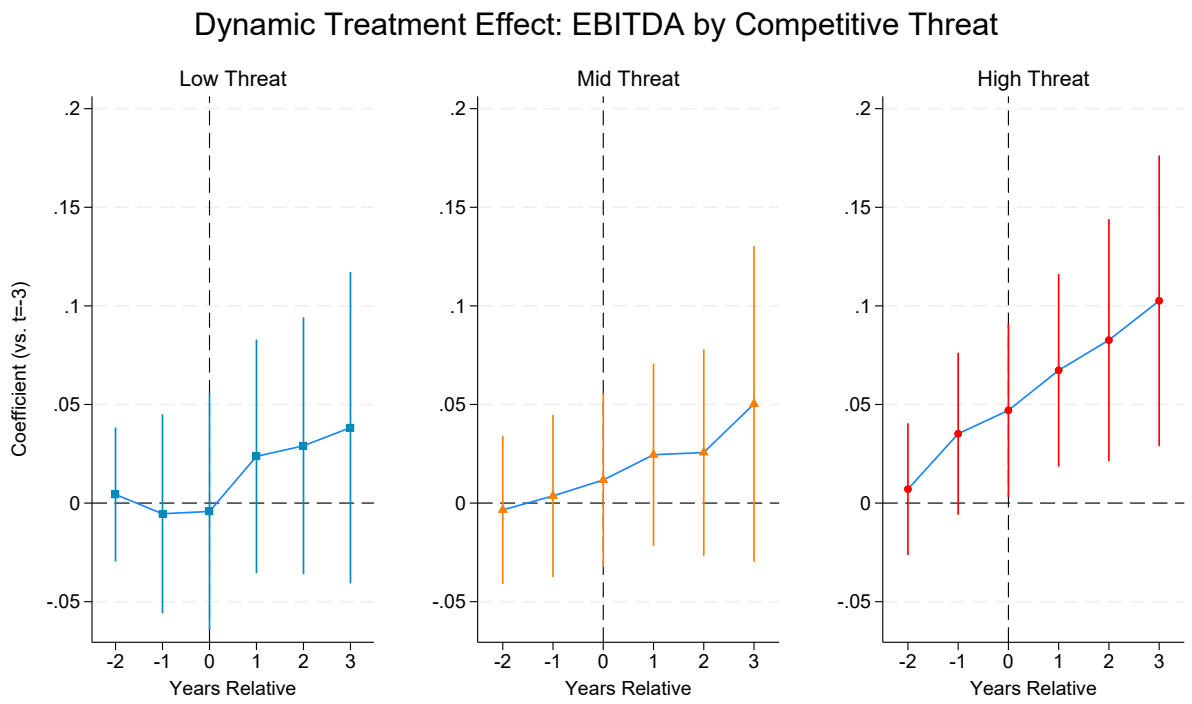



**Figure IA2. Dynamic Treatment Effect on EBITDA by Competitive Threat**

This figure presents the dynamic treatment effects on the profitability (EBITDA) of peer firms across three competitive threat groups (Low, Mid, and High Threat). The x-axis indicates the years relative to the event year (Year 0), and the y-axis shows the estimated coefficients representing the difference from the base year ($t = -3$), with corresponding 95% confidence intervals.

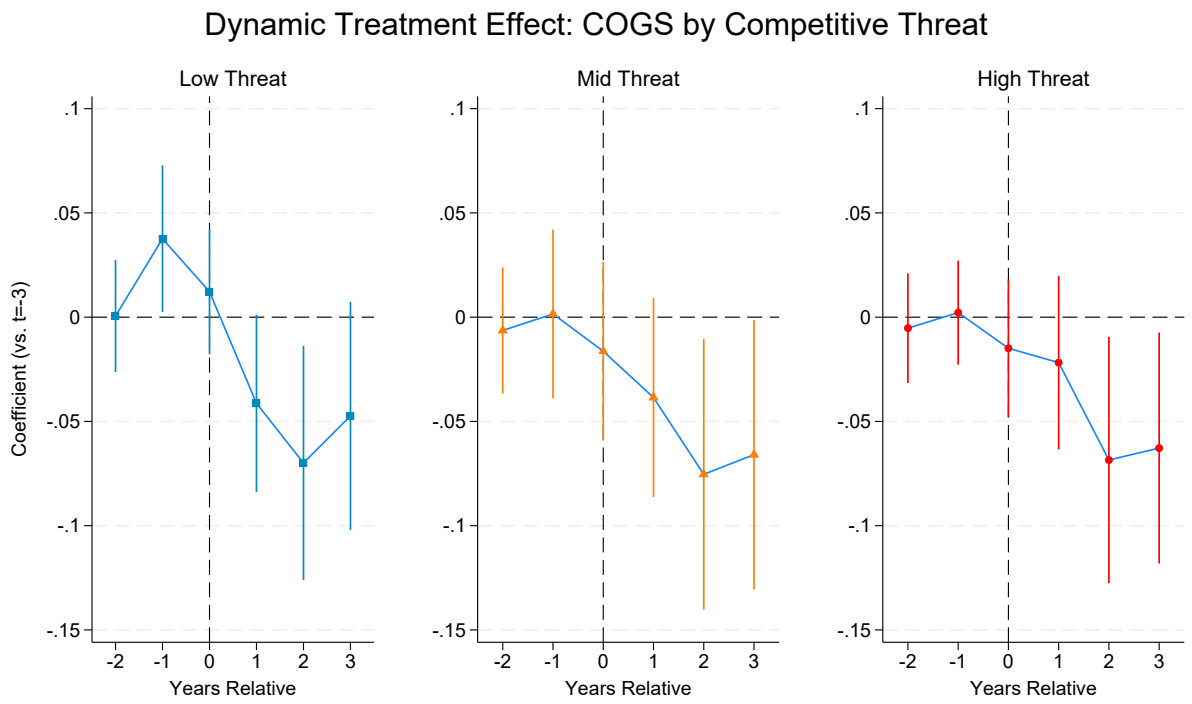



**Figure IA3. Dynamic Treatment Effect on COGS by Competitive Threat**

This figure illustrates the dynamic treatment effects on the Cost of Goods Sold (COGS) for peer firms, categorized by their initial competitive threat level (Low, Mid, and High Threat). The x-axis denotes the years relative to the open-source release (Year 0). The y-axis represents the estimated coefficients relative to the base year ($t = -3$), plotted with their 95% confidence intervals.

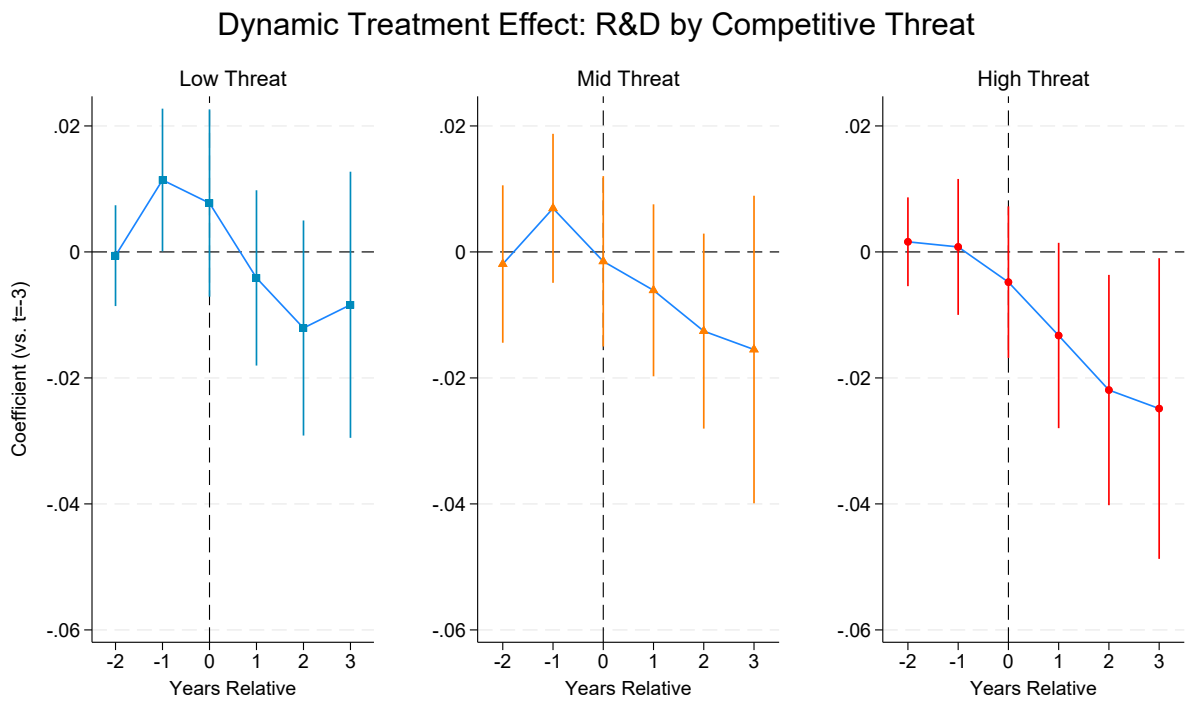



**Figure IA4. Dynamic Treatment Effect on R&D by Competitive Threat**

This figure plots the dynamic treatment effects on the Research and Development (R&D) expenditures of peer firms across Low, Mid, and High threat groups. The x-axis represents the years relative to the focal firm's open-source release (Year 0). The y-axis displays the estimated coefficients as the difference from the base year ($t = -3$), along with their 95% confidence intervals.

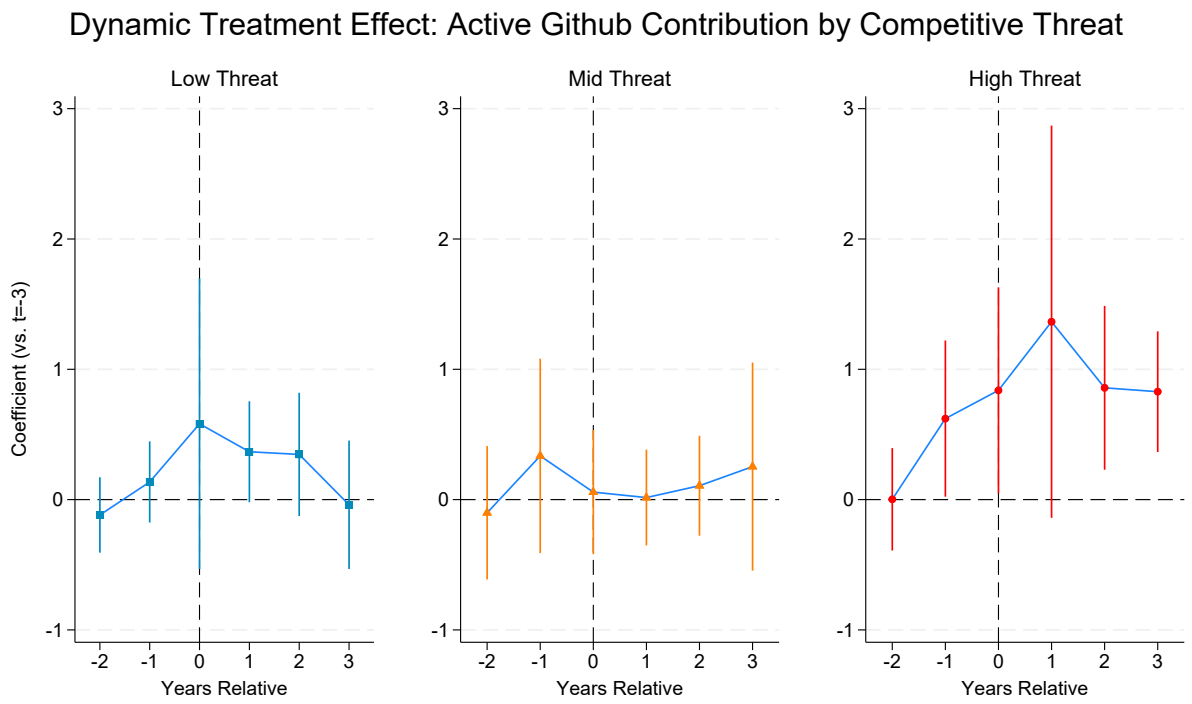



**Figure IA5. Dynamic Treatment Effect on Active GitHub Contribution by Competitive Threat**

This figure displays the dynamic treatment effects on the active GitHub contributions of peer firms, stratified by their initial competitive threat level (Low, Mid, and High Threat). Active contributions include direct development efforts such as pushes, pull requests, and issues. The x-axis indicates the years relative to the open-source release (Year 0), and the y-axis shows the estimated coefficients relative to the base year ($t = -3$) with 95% confidence intervals.

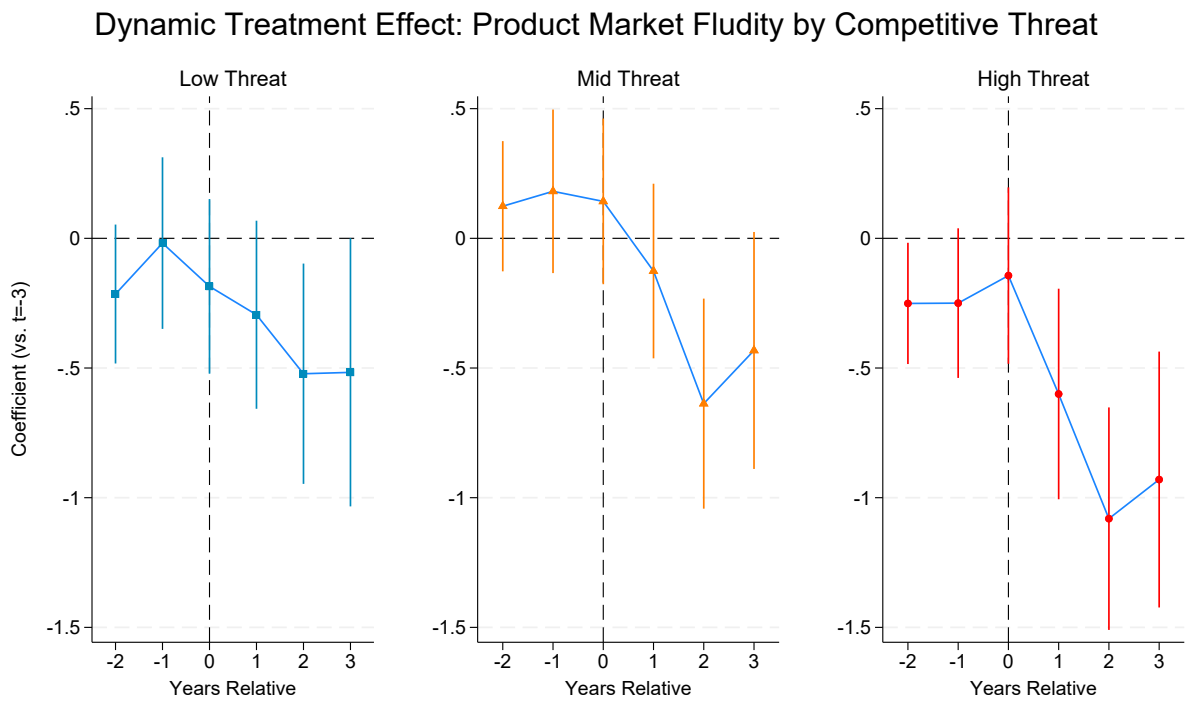



**Figure IA6. Dynamic Treatment Effect on Product Market Fluidity by Competitive Threat**
This figure plots the dynamic treatment effects on the product market fluidity of peer firms across three competitive threat groups (Low, Mid, and High Threat). Product market fluidity measures the degree of competitive threat based on changes in product descriptions. The x-axis represents the years relative to the focal firm's open-source release (Year 0). The y-axis displays the estimated coefficients as the difference from the base year ($t = -3$), with 95% confidence intervals included for each point estimate.