Commoditize Your Complement:
Who Bears the Cost of an Open-Source Release?

Peter Young Won *

November 2026

ABSTRACT

Why do large technology firms give away open-source software for free, and who pays for it? I argue that complement firms pay for it, not substitutes, and that the release is therefore a strategy to commoditize its complements. To identify complement and substitute peers, I place each peer on two axes: technological proximity to the released software and product-market overlap with the releasing firm. Across 249 releases, I find that complements lose 0.73% at the release and substitutes do not react, whereas the releasing firm earns 1.54% over the following month. I further document that complement peers lose sales, cut investment and R&D, obtain 21% fewer patents, and more than double their active open-source contributions. Alternative explanations such as distress and cost cutting do not account for a decline confined to complement peers.

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

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

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

1 Introduction

Whether the technology that one firm creates helps or harms other firms in the same industry is a central question in the economics of innovation (Jaffe, 1986; Bloom et al., 2013). On the one hand, a new technology helps the firms that adopt the technology, because an adopter obtains a capability without paying to build the capability. On the other hand, the same technology harms the firms that sell a functional equivalent of the technology, because a cheaper product replaces what those firms had been selling. Whether a new technology benefits or harms a given firm therefore depends on how that firm stands in relation to the technology and to the firm that created the technology (Cournot, 1838; Varian, 2004). New technologies arrive continually, so that relation has to be measured systematically and on economic grounds rather than case by case.

Corporate open-source software is a setting in which a single act has both effects. Alphabet published Kubernetes and the Go programming language, Meta published PyTorch, and Microsoft published the compiler for TypeScript. Each release is free of charge and available to every firm in the industry. Giving away a finished product is an unusual competitive act, because the releasing firm earns no revenue on the product and cannot withdraw the product later. Who gains and who loses from such a release is not obvious.[1] Recent work measures the gains at adopting firms (Nagle, 2019; Haese and Peukert, 2025), at publishing firms (Emery et al., 2025; Coleman et al., 2025; Conti et al., 2025), and in the economy as a whole (Hoffmann et al., 2024), but the losses have not been measured systematically. Therefore, this paper asks which firms bear the cost when a large technology firm gives away open-source software.

For each release, I classify the peer firms closest to the released software into two groups. Complements sell what the released software replaces, and sell outside the product market of the releasing firm. Substitutes are as close to the released software as complements, and sell inside that market. I find that an open-source release reduces the value of complements and does not affect substitutes. The equity market separates the two groups within a day of the release. Over the following three years, complements sell less, invest less, employ fewer workers per dollar of sales, and are granted fewer patents. Substitutes change on none of those margins. Active contributions

[1] Writing outside the academic literature, Gurley (2026) observes that the owners of the cloud stack, IBM, Hewlett-Packard, and Intel, did not capture cloud computing, because open-source commoditization dissolved the supplier power those firms had held. In that episode commoditization was nobody's strategy. The releases studied here are deliberate.

to open-source repositories more than double at complements and rise by a fifth at substitutes. Financial distress and cost cutting do not account for a decline confined to complements. These results indicate that an open-source release is competitive conduct rather than a benefit to the industry at large. A release commoditizes the complements of the releasing firm, and not the firms that compete with the releasing firm.

These results are not obvious ex ante, because two strands of the literature make opposite predictions about the firms around a release. The first strand treats an open-source release as a public good that raises the value of every firm able to use it (Nagle, 2019; Hoffmann et al., 2024). The second strand treats the same release as competitive conduct that reduces the value of the firms selling the commoditized product (Cournot, 1838; Varian, 2004). Because the two predictions apply to different firms, the average effect across the firms around a release cannot tell them apart, and only the sign of the effect within each group reveals which prediction holds.

I posit that the cost of a release falls on the firms that sell what the released software replaces, and not on the firms that compete with the releasing firm. The reason is that a release makes a complement to the releasing firm's core business free (Cournot, 1838; Parker and Van Alstyne, 2005). A firm that sells the commoditized product cannot sustain a positive price once a functional equivalent is available at no cost. A firm that competes with the releasing firm sells a product for which the released software is not a substitute, so the price of that product is unaffected. Therefore, the two groups are exposed to the same release with opposite sign.

Identifying which firms are exposed is challenging because a release is not a shock to a firm but a shock to a position. Consider a firm that sells a proprietary machine-learning framework and a firm that sells search advertising. Both firms are technology firms, both appear as product-market competitors of Alphabet, and both work on similar problems. A release by Alphabet nonetheless reaches the two firms with opposite sign. A design that treats every product-market competitor as equally exposed averages the two firms and recovers an effect near zero. Bloom et al. (2013) face the same problem for research and development and address the problem by placing every pair of firms on a technology axis and a product-market axis at once. I apply that decomposition to a shock that can be measured directly. The technology axis scores the released repository against the goods and services the peer registers as trademarks, and the product-market axis uses the peer's segment sales and trademarked goods.

To apply the decomposition, I screen the event set on what the released software is rather than on how widely the software came to be used. Listed firms own 151,085 public repositories, most of which are tutorials, sample code, and small libraries. I keep the 321 repositories that are finished products a third party can build, run, and ship in place of software the third party would otherwise have purchased. Stars, forks, and downloads never enter the screens, because those measures accumulate over the years after a release and are correlated with the outcomes estimated here. Of the 321 releases, the 249 releases by 41 firms that have the peer pools and the data both estimators require form the sample.

To separate the two exposed groups, the classification uses the two axes for different purposes. Complement and substitute peers have mean technology ranks of 0.80 and 0.86 and mean product-market ranks of 0.18 and 0.85. The two groups are therefore about equally close to the released software and far apart in where they sell. The technology axis changes from one release of a firm to the next and therefore determines whether a peer is exposed at all. A complement peer of one release is a control peer of another in 95 percent of cases. The product-market position of a peer relative to a given releasing firm changes little and therefore determines which side an exposed peer falls on. The label switches between complement and substitute in 0.1 percent of cases.

I first examine whether the equity market prices a release differently across the two groups. I measure the abnormal return of each group against the Carhart (1997) four factors. The abnormal return is estimated on calendar-time portfolios rather than averaged across events, because the windows around the releases of one firm overlap. I find that the complement portfolio earns -0.73 percent over the four trading days that begin with the release, and that the substitute portfolio does not move in any window. Neither portfolio moves in the ten trading days before the release. The releasing firm does not move at the release and earns 1.54 percent over the following month. These results suggest that the market treats a release as bad news for the firms that sell what the released software replaces and as no news for the firms that sell where the releasing firm sells. The releasing firm gains over the month that follows, so a release is not priced as a cost to the industry as a whole.

I then ask whether the repricing corresponds to a change in what the exposed firms sell and spend. I estimate a stacked difference-in-differences at annual frequency on the same 249 releases, over event years -3 to $+3$. Each release defines a cohort, and the coefficients compare exposed peers with control peers of the same release in the same fiscal year. Control peers are reweighted

by entropy balancing to match the exposed group on pre-release levels and trends (Hainmueller, 2012). I find that complement peers lose 5 percent of sales, cut capital expenditure by 12 percent, and employ fewer workers per dollar of sales. Substitute peers, compared with control peers matched to the substitute group, change on none of the three margins. These results suggest that the loss is concentrated on complement peers and appears in operations rather than in expectations alone. The pattern is consistent with a release that commoditizes a complement of the releasing firm's core business (Cournot, 1838; Parker and Van Alstyne, 2005).

I further examine how complement peers develop software after a release. I find that complement peers cut research and development by 4 percent of the pre-release level and are granted 21 percent fewer patents, including 17 percent fewer patents in artificial intelligence. Over the same years, active contributions by complement peers to open-source repositories rise by 113 percent, against 21 percent at substitute peers. Active contributions are forks, issues, comments, pull requests and pushes, and exclude watch events. Contributions to a public repository are timestamped events attributable to the contributing firm, so the reallocation of engineering effort is measured directly rather than inferred from R&D expense. These results suggest that complement peers respond to a release by moving engineering effort from proprietary development toward open-source code.

One potential concern is that financial distress or a restructuring unrelated to the release explains the decline at complement peers. I therefore examine the balance sheet and the cost structure of complement peers after a release. Complement peers reduce debt, add cash, and show no rise in default risk on either main measure, which is not the balance sheet of a firm in distress. Operating income at complement peers falls while overhead rises relative to sales, which is not the pattern of a firm cutting costs.

This paper contributes to three strands of literature. First, the paper adds to the empirical literature on corporate open source. That literature measures what the releasing firm or the adopting firm captures (Nagle, 2019; Hoffmann et al., 2024; Emery et al., 2025; Coleman et al., 2025; Haese and Peukert, 2025; Zhao, 2025; Nagaraj and Piezunka, 2024; Conti et al., 2025), and treats a release as an act that creates value. I show that a release also transfers value, and that the firm selling what the released software replaces bears the cost. The firm being commoditized is the unit of analysis here.

Second, the paper contributes to the literature on how the technology of one firm reaches another firm. A large literature separates the knowledge that one firm's research provides to others from the rivalry the research creates, and finds that spillovers dominate on average (Jaffe, 1986; Bloom et al., 2013; Aghion et al., 2005; Makaew and Maksimovic, 2020; Baker et al., 2023). Similarity between firms is the standard way of locating exposure, and similarity cannot make the distinction this setting requires. Complement peers are 14 percentage points less likely than substitute peers to be listed as product-market rivals of the releasing firm by Hoberg and Phillips (2016). A measure of rivalry would therefore identify the group that loses value as the less exposed group. I show that exposure to a release can be measured from the released artifact itself, which makes exposure a property of an event rather than of a pair of firms.

Third, the paper contributes to the literature on platform owners and the firms that build on a platform. That literature studies a platform owner entering the market of a complementor (Eisenmann et al., 2011; Boudreau, 2010; Seamans and Zhu, 2014; Wen et al., 2016; Cennamo et al., 2018; Wen and Zhu, 2019). An open-source release commoditizes a market without entering the market. I show that a release compresses the revenue of the firms selling in the commoditized layer, does not affect the firms competing with the releasing firm, and changes no price the releasing firm charges. Whether such conduct warrants regulatory attention depends on effects on consumers this design does not measure.

2 Institutional Background and Theoretical Predictions

This section describes open-source software and the platform on which it is published, the terms under which a firm releases source code, and the economics that determine which firms a release reaches.

2.1 Open-Source Software and GitHub

Open-source software is software whose source code is published under a licence that allows anyone to use, study, modify, and redistribute it.[2] Source code is the human-readable text from

[2]The Open Source Initiative maintains the Open Source Definition that most such licences follow, `https://opensource.org/osd`.

which a program is built, so a firm that obtains the source code can compile the program, change it, and ship it inside a product of its own. Proprietary software is instead sold as a compiled program under a licence that forbids modification and redistribution. Because anyone may copy and redistribute an open-source program, the market price of the program is zero. Licences differ in what they require of the recipient. Permissive licences, such as the MIT and Apache licences, allow the code to enter a proprietary product. Copyleft licences, such as the GNU General Public License, require a recipient that distributes a modified program to publish the modified source code.

Under both types of licence, the grant to recipients cannot be withdrawn.[3] A firm can change the licence of future versions, but earlier versions remain open. In January 2021, Elastic moved new versions of its Elasticsearch search engine to licences that the Open Source Initiative does not recognize as open source, and in April 2021 Amazon released OpenSearch, built from the last open version of Elasticsearch.[4] OpenSearch is one of the releases in the sample. Accordingly, a firm that publishes a program cannot later restore the price that the release set to zero.

Open-source software differs from the open-weight models that artificial intelligence firms have released since 2023. An open-weight release publishes the trained parameters of a model, which allows a recipient to run and fine-tune the model but not to rebuild it, because the training data and the training pipeline stay with the releasing firm.[5] The releases studied here are code repositories, so the code published alongside a model can enter the sample while the weights cannot.

Open-source software is a general input to production rather than a niche. Nagle (2019) finds that firms using open-source software are more productive than firms that do not. Hoffmann et al. (2024) estimate the demand-side value of widely used open-source software at $8.8 trillion and

[3]The Apache License, Version 2.0, grants every recipient a perpetual, worldwide, royalty-free, and irrevocable copyright licence, `https://www.apache.org/licenses/LICENSE-2.0`.

[4]Elastic announced the change on 14 January 2021, `https://www.elastic.co/blog/licensing-change`. The Open Source Initiative explains why the Server Side Public License is not an open-source licence at `https://opensource.org/node/1099`. Amazon announced OpenSearch, derived from Elasticsearch 7.10.2, on 12 April 2021, `https://aws.amazon.com/blogs/opensource/introducing-opensearch/`.

[5]In October 2024, the Open Source Initiative published version 1.0 of its definition of open-source artificial intelligence, which requires detailed information about the training data in addition to the complete code and the parameters, `https://opensource.org/ai/open-source-ai-definition`. Open-weight licences also often restrict use. The licence for Meta's Llama 2 models, for example, required a licensee whose products had more than 700 million monthly active users to request a separate licence from Meta, `https://github.com/meta-llama/llama/blob/main/LICENSE`.

find that firms would need to spend 3.5 times more on software if open-source software did not exist. Because the price of the program is zero, neither the value a release creates nor the revenue it displaces appears in a market price.

Open source began with volunteer communities, which built projects such as the Linux operating system and the Apache web server (Lerner and Tirole, 2002). Large technology firms now publish at scale. The organizational accounts of 591 publicly traded US firms own 151,085 public repositories, and the 249 releases studied here are dated from 2010 to 2024. Publication is skewed across these firms rather than confined to a few of them. The four largest publishers, Microsoft, Amazon, Alphabet, and IBM, own 32 percent of the repositories, and 15 firms own half, while the median firm owns 52 repositories and the mean firm owns 256. Firms also contribute to repositories they do not own, and far more firms take part in open source than publish their own code. GitHub accounts are attributed to 4,467 firms, against the 591 firms that own public repositories.

Most open-source software is developed and published on GitHub, a platform that hosts code repositories and records every change made to them. GitHub reported more than 100 million developers in January 2023, and Microsoft acquired GitHub in 2018 for $7.5 billion in Microsoft stock.[6] A repository holds the source code of one project together with the history of its changes. A developer writes code locally and pushes a commit, which is a recorded change, to the repository. A developer without write access forks the repository, which copies the repository under that developer's own account, and then opens a pull request asking the owner to merge the change back. An issue is a public report of a defect or a request for a feature, and the comments on an issue are public as well. A watch event records a user following a repository without contributing to it. Every such event includes an account identifier and a timestamp, which is what allows the contributions of a firm's engineers to be counted.

2.2 Theoretical Predictions and Hypotheses

The price of open-source software is zero, which distinguishes an open-source release from the ordinary introduction of a product. The zero price binds on both sides of the market. A firm that publishes source code earns nothing on the program and cannot later charge for it, and every firm

[6] `https://github.blog/2023-01-25-100-million-developers-and-counting/` and `https://news.microsoft.com/2018/06/04/microsoft-to-acquire-github-for-7-5-billion/`.

in the industry obtains the program on the same terms, including the firms that had been selling such a program. Publishing a program is in that sense costly to the releasing firm as well, and a firm accepts that cost only where the release does not reduce its own revenue.

What a release costs the publisher depends on the market the release enters, so the market a firm publishes into is not arbitrary. A program that replaced the firm's own core product would cannibalize the rents the firm earns on that product. A firm therefore publishes a program whose cheapening raises demand for the product the firm does sell. Two goods standing in that relation are complements, and two goods for which a fall in the price of one lowers demand for the other are substitutes (Cournot, 1838). An open-source release therefore enters the market for a complement of the releasing firm's product rather than the market the releasing firm serves, and the incidence of the release follows from where the release lands. A peer selling in that market faces a functional equivalent of its own product at no cost, under a licence that permits redistribution, and cannot sustain a positive price. A peer that competes with the releasing firm, by contrast, sells a product the release does not replace, so the price of that product is unchanged, and that peer obtains the program on the same terms as any firm. One open-source release therefore reaches the two groups of peers with opposite sign.

Those two relations map onto the vertical organization of software production. An application runs on a runtime, the runtime runs on an operating system, and the whole stack runs on hardware. Firms at adjacent layers sell to the same customer without selling the same product, whereas firms at one layer sell that customer the same product. Adjacent layers therefore hold complements, because a customer who pays less for a runtime buys more of the hardware beneath it, and a single layer holds substitutes. Because the relation is a property of the goods and not of the firms that supply them, two firms with similar technology can fall on opposite sides of it. An open-source release therefore reprices a position in the stack rather than a firm, and the firms that occupy that position differ from one release to the next.

This mechanism differs from the channel through which the technology of one firm ordinarily reaches another. A technology spillover transmits knowledge, which raises the productivity of a firm able to absorb the knowledge (Jaffe, 1986). An open-source release instead transmits a finished, licensed, runnable program, which changes the price a peer can charge for a product the peer already sells.

The argument above is not the only prediction the literature supports, and two strands of the literature make opposite predictions about the firms around an open-source release. The first strand treats a release as a public good. Adoption of open-source software raises productivity at the firms that use the software (Nagle, 2019), and the value created accrues broadly rather than to the releasing firm alone (Hoffmann et al., 2024). A public good therefore raises the value of every firm that can use it. The second strand treats a release as the competitive conduct described above (Cournot, 1838; Varian, 2004; Rochet and Tirole, 2003; Parker and Van Alstyne, 2005; Armstrong, 2006; Economides and Katsamakas, 2006; Gawer and Henderson, 2007). Zero marginal cost, network benefits, and standardization also make a price of zero consistent with equilibrium for a digital good (Katz and Shapiro, 1985; Farrell and Saloner, 1985; Shapiro and Varian, 1999; Casadesus-Masanell and Ghemawat, 2006). Competitive conduct therefore reduces the value of the firms that sell the commoditized product. Because the two strands predict losses at different firms, the average response of the firms around a release cannot separate them.

Both strands concern the effect of an open-source release on other firms rather than the reason a firm publishes. A firm that earns on a product the released software makes more valuable has a reason to publish, and Spolsky (2002) describes that practice as commoditizing a complement.[7] This paper measures where the cost of an open-source release falls and not why the release was made, because a release reduces the price of a complement whether or not the reduction is the reason the firm published the code.

Each of the two strands implies a hypothesis about which peers lose value, and a third hypothesis is that an open-source release is entry into the releasing firm's own market. Under the public-good hypothesis, peers close to the released software gain whether or not they sell where the releasing firm sells. Under the entry hypothesis, the peers that sell where the releasing firm sells lose and the peers that sell elsewhere do not. Under the commoditization hypothesis, the peers that sell what the released software replaces lose and the peers that compete with the releasing firm do not. The commoditization hypothesis is the one this paper tests, and it assigns opposite signs to the two exposed groups.

[7] Firms publish source code for other reasons as well. Released code can attract outside developers who improve it, raise the firm's standing among engineers, and set a standard that other firms build on (Lerner and Tirole, 2002; Henkel, 2006; Fosfuri et al., 2008; Alexy et al., 2013; Casadesus-Masanell and Llanes, 2011), and the resource-based view explains when a firm gains by giving up control of a resource it owns (Barney, 1991; Chesbrough and Crowther, 2006; Alexy et al., 2018). Section 5.2 asks whether those motives predict the pattern documented here.

Hypothesis 1. *If an open-source release commoditizes a complement of the releasing firm's core business, the released software reaches the peers that sell that complement as a free rival. Those peers lose value at the release and contract over the years that follow.*

Hypothesis 2. *Under the same condition, the released software reaches the peers that sell where the releasing firm sells as a free input. Those peers do not lose value and do not contract, although they are as close to the released software as complement peers.*

The two hypotheses concern the sign within each group rather than the average across the peers of a release, so testing them requires a classification that separates the groups before any outcome is observed.

Three features of corporate open source make the hypotheses testable. First, the date of an open-source release is public and precise, because GitHub records the moment a repository becomes public. Second, an open-source release that commoditizes a product is identifiable, because most of what a firm publishes is documentation, sample code, or a library rather than a finished program. Third, an open-source release is free to every firm and cannot be withdrawn. Overall, an open-source release of a finished program is a dated, permanent, and zero-priced entry into the market for the product the program replaces.

3 Data and Measurement

Measuring who bears the cost of an open-source release imposes two requirements on the design. First, membership in the event set must be determined without reference to how the market or the peers responded after a release, so that selection into the sample cannot be driven by the outcomes under study. Second, the classification of peer firms must vary across the releases of a single firm, because the shock is an individual program rather than the firm that published it. This section describes the release sample, the classification of peer firms, and the outcome variables.

3.1 The Release Sample

I begin with the 151,085 public GitHub repositories owned by the organizational accounts of 591 publicly traded firms.[8] The event is the day a firm makes the source code of a repository public. Table 1 reports every screen in the order it is applied, together with the number of repositories each screen removes.

[Insert Table 1 here]

Most repositories on a firm's public account are not software that a third party could use in place of a product. The accounts contain sample code, tutorials, documentation, configuration files, mirrors, and clients for services the firm sells. Accordingly, the first screens remove such repositories mechanically. The mechanical screens keep repositories that have a recorded release date, are not forks or templates, and have a description, a detected programming language, and a software licence. The mechanical screens also require that the owner was listed with US common shares when the repository became public. Screens on the repository name, language, and topics then remove teaching material, documentation, mirrors, personal repositories, clients for the firm's own services, notebooks, and duplicates, which retains 22,265 repositories.

The last two screens are manual, and each remaining repository is read against criteria fixed in advance. The first manual screen removes teaching material, asset dumps, and mirrors that the mechanical screens missed. The second manual screen keeps a repository only if it is a finished artifact that the firm presents as a product in its own right. Platforms, runtimes, programming languages, engines, databases, and applications pass the screen. Libraries, plugins, extensions, and components do not pass, because they add a feature to a product rather than replace the product. The two screens retain 321 releases.

No screen measures how widely the software came to be used. Stars, forks, and downloads accumulate over the years after a release and are correlated with the outcomes estimated here, so selecting releases on those measures would introduce selection bias.

The date of a release is the moment the repository became public on GitHub, or the repository's creation when GitHub records no public event. Not every dated release can be estimated, because

[8] Documentation for the GitHub REST API is at `https://docs.github.com/en/rest`. Event-level activity comes from the GH Archive public dataset on Google BigQuery, `https://www.gharchive.org/`.

a release is informative only when the industry around it contains both an exposed peer and a peer to compare that peer with. A release therefore enters the estimation sample when the computer, semiconductor, and software firms matched to the release include at least one complement peer and one control peer with return and accounting data. The estimation sample comprises 249 releases from 41 firms from 2010 to 2024, and every exhibit in this paper is estimated on this sample. Table IA.1 lists all 249 releases. Figure 1 plots the releases by year and by releasing firm. Panel (a) shows that releases are rare before 2014 and then run between 14 and 41 a year, with 90 percent of the sample dated 2014 or later, and the largest number in 2018. Panel (b) shows that the eight firms with the most releases account for 195 of the 249 releases, and that the remaining 33 firms account for 54.

[Insert Figure 1 here]

The screens can be checked against the software they retain. Table 2 lists thirty of them, including Alphabet's Kubernetes and Go, Microsoft's TypeScript and ONNX Runtime, Meta's PyTorch, Oracle's MySQL and GraalVM, Intel's OpenVINO, and Nvidia's TensorRT-LLM. Each release is software that a firm outside the releasing company can build into a product of its own, which is the property the second manual screen tests for.

[Insert Table 2 here]

The sample is concentrated in a few releasing firms. Microsoft accounts for 70 releases, Alphabet for 40, Oracle for 27, and Meta for 19, and Table IA.2 lists all 41 releasing firms. Table 3 describes the repositories and the releasing firms. The median repository was built privately for 48 days before its release and is matched to 319 peers. The median releasing firm-year has $99.8 billion of assets and spends $6.0 billion on R&D. The concentration follows the product screen rather than publication volume. Amazon and IBM own the second and fourth largest repository counts in the universe but rank fifth and tenth by releases, whereas Oracle ranks 18th by repository count and third by releases. The concentration nonetheless raises a concern about external validity. I acknowledge that the estimates describe large releasing firms and do not extend to smaller ones. However, the question itself concerns firms of this kind. A firm gains from commoditizing a complement only when the firm earns rents in a core business whose demand a cheaper complement

raises (Cournot, 1838). Re-estimating the main specifications without each releasing firm in turn shows that no single releasing firm drives the estimates.[9]

[Insert Table 3 here]

Overall, the release sample is built from what the software is and not from how the software was received. The screens reduce the 151,085 repositories of 591 listed firms to the 22,265 that survive the mechanical stage. A manual reading then finds 321 of those to be finished products, and 249 of the 321 have the peer pools and the data the estimators require. No screen uses a return, an accounting outcome, or a measure of adoption, so membership in the event set is fixed before any response to a release is observed.

3.2 A Release-Level Classification of Peer Firms

A single measure of how related two firms are cannot separate a release that helps a peer from one that harms a peer. To separate the two, I place each peer on a technology axis and a product-market axis at once, following Bloom et al. (2013), and I build both axes around the released repository rather than around the releasing firm.

Both axes rest on the uncentered correlation used by Jaffe (1986) and Bloom et al. (2013). For two share vectors S_a and S_b,

$$\text{PROX}_{ab} \;=\; \frac{S_a S_b'}{\sqrt{S_a S_a'}\,\sqrt{S_b S_b'}}, \tag{1}$$

which equals one when the two vectors have the same proportions and zero when they share no element.

The peer pool of a release consists of the public computer, semiconductor, and software firms in SIC codes 3570 to 3579, 3674, and 7370 to 7379 that have the data both axes require. The median release has 319 peers. A firm that reports the releasing firm as a customer in the Compustat customer segment data is dropped from the pool of the releasing firm. A supplier is affected by its customer's release through the sales it makes to the customer, which is a different relation from the two the classification measures (Adner and Kapoor, 2010).

[9]Each releasing firm with at least five releases is dropped in turn, and every subsample preserves the sign and significance of the complement return over $[0,+1]$ and of every panel coefficient.

The technology axis measures whether the released software is a functional equivalent of the peer's products. The natural candidate, following Jaffe (1986) and Bloom et al. (2013), is a profile of patent classes. A patent-class profile is a property of the firm, however, so it takes the same value for every repository the firm publishes, while the technology axis has to vary with the repository released. I therefore measure the technology axis with trademarks, which are registered for the goods and services a firm sells rather than for the firm as a whole (Mendonça et al., 2004; Graham et al., 2013; Heath and Mace, 2020). Every registered trademark has a filing date and a description of the goods and services it covers, written to pass examination at the US Patent and Trademark Office. Because only marks filed on or before the release year enter, the technology axis uses no information from after the release.

The technology axis combines two measures. First, the goods-text similarity is the text similarity between the repository and the goods-and-services descriptions of the peer's marks, and it measures how close the peer's products are to the software.[10] Second, the live-mark share is the share of the peer's live marks whose descriptions are among the five percent of live marks most similar to the repository. I rank both measures within the peer pool of the releasing firm, and the technology score is 0.67 times the rank of the goods-text similarity plus 0.33 times the rank of the live-mark share.

The product-market axis measures whether the peer sells where the release competes, and it also combines two measures, each ranked within the release. First, the segment overlap applies equation (1) to Compustat segment sales. I locate the release in segment space by averaging the segment-sales shares of the fifteen firms whose 10-K business descriptions are most similar to the repository, weighted by the similarity and excluding releasing firms. The segment overlap is the correlation between the resulting vector and the peer's segment-sales shares.[11] Second, the trademark overlap applies equation (1) to the goods-and-services vectors of the peer's marks and the releasing firm's marks.

Crossing the two axes yields three groups. A complement peer is in the top third of the technology axis and in the bottom third of both product-market measures. The industry of a complement

[10] The repository text is its name, description, topics, and primary language. The similarity is the cosine between term vectors weighted by inverse document frequency.

[11] Segment-sales shares are pooled over the sample period, as in Bloom et al. (2013), so the segment overlap uses sales reported after the release. The trademark measures use only marks filed on or before the release year.

peer must in addition trade with the releasing firm's industry, in either direction, in the 2017 detail Use table of the Bureau of Economic Analysis. A substitute peer is in the top third of the technology axis and in the top third of both product-market measures. A control peer is at or below the median of the technology axis and at or below the 75th percentile of both product-market measures. Requiring the two product-market measures to agree keeps the exposed groups small, and a release needs at least nine peers.

The three conditions on a complement peer state that the software is a functional equivalent of the peer's products, that the peer does not sell where the release competes, and that the two industries buy from one another. The first two place the peer in the corner that is close in technology and distant in product market, which is the position Section 2.2 identifies as the one a release reprices. The third asks for a positive relation rather than the absence of rivalry, because a firm that merely fails to compete with the releasing firm is not thereby a complement of it. The split between the two exposed groups follows the distinction between firms that supply a complement and firms that supply a substitute (Adner and Kapoor, 2010; Adner and Lieberman, 2021).

The three groups discretize the two axes of Bloom et al. (2013), who place each pair of firms in a unit box and discuss the corner that is close in technology and distant in product market. Bloom et al. (2013) enter the two measures as continuous weighted sums of the research of other firms, because they estimate an elasticity of performance to that research. This paper compares how groups of peers respond to one event, so the two measures enter as groups. Adopting their structure does not mean adopting their prediction, because what crosses the boundary here is a finished program rather than knowledge, as Section 2.2 explains.

Cutting a continuous proximity measure at a percentile is how the literature on economically linked firms identifies complementors. Lee et al. (2024) keep peers in the top two percent of a production-complementarity score, drop peers in the same four-digit industry and peers with a material customer-supplier relationship, and require at least three complementor peers for a firm. Those three choices correspond to the cut at the extreme of the distribution, the filter for vertical relations, and the floor on the number of peers imposed here. However, Lee et al. (2024) identify complementors alone, whereas the product-market axis here also separates complement peers from substitute peers.

Table 4 reports the classification. In Panel A, complement and substitute peers have mean

technology ranks of 0.80 and 0.86 and mean product-market ranks of 0.18 and 0.85. This implies that the two exposed groups are about equally close to the software but far apart in where they sell. The estimation sample contains 890 complement, 5,065 substitute, and 14,388 control pairs of a release and a peer with pre-release data, drawn from 218, 277, and 414 distinct firms. Table IA.3 names the complement and substitute peers of every release.

[Insert Table 4 here]

Panel B of Table 4 compares the three groups before the release. Complement peers are smaller than control peers by 1.252 log points of assets ($t = -3.19$) and smaller than substitute peers by 1.368 ($t = -4.27$). Their sales relative to assets exceed those of control peers by 0.177 ($t = 2.55$) and those of substitute peers by 0.208 ($t = 3.32$). Market-to-book, leverage, R&D, and investment do not differ between complement peers and either other group. This implies that complement peers are smaller and sell more relative to their assets than the peers they are compared with, while the remaining characteristics are balanced.

Each axis is a combination of two measures, so the separation between the groups could rest on one measure alone. Figure 2 plots all three groups on the two axes and on the four measures that enter them. Control peers have mean technology and product-market ranks of 0.24 and 0.35, so they are far from the released software on the axis that decides exposure. The median goods-text similarity is 0.044 for complement peers, 0.058 for substitute peers, and 0.016 for control peers, and the median live-mark share is 0.125, 0.097, and 0.000. The median segment overlap is 0.134 for complement peers, 0.726 for substitute peers, and 0.344 for control peers, and the median trademark overlap is 0.094, 0.332, and 0.126. Each measure separates the three groups in the same direction as the axis it enters.

[Insert Figure 2 here]

The classification varies across the releases of a single firm. Consider the 548 pairs of a releasing firm and a peer in which the peer is a complement peer for one release and is classified for another release of the same firm. In 95 percent of the pairs, the peer is a control peer for another release, whereas a firm-level rule would assign each peer the same label for every release. Labels rarely switch between the two exposed groups, however. Among peers that a releasing firm places

in an exposed group more than once, the label switches between complement and substitute for 0.1 percent. This implies that the technology axis changes from release to release and determines whether a peer is exposed, while the product-market position of a peer relative to a given releasing firm changes little.

To test whether the product-market axis measures where peers sell, I examine measures that the classification never uses. If the product-market axis captures where peers sell, substitute peers should be closer rivals of the releasing firm than complement peers on outside measures of rivalry. If the product-market axis instead captures noise, outside measures of rivalry should not separate the two groups. Table 5 regresses three such measures on an indicator for complement rather than substitute peers, with release and peer-firm fixed effects. The three measures come from sources that share nothing with one another, namely the product descriptions in 10-K filings, an industry code, and an accounting series. The coefficient on Complement is -17.76 for being listed as a product-market rival of the releasing firm in the TNIC data of Hoberg and Phillips (2016) ($t = -5.34$), against a substitute-peer mean of 26.49 percent. A substitute peer is therefore nearly five times as likely to be a listed rival as a complement peer.

The remaining two columns use data other than text. Complement peers are 37.17 percentage points less likely to report the releasing firm's own four-digit SIC code ($t = -5.76$), against a substitute-peer mean of 35.00 percent. The annual sales growth of complement peers co-moves less with that of the releasing firm over the six fiscal years before the release, by 0.2032 of a correlation ($t = -2.35$). The mean is -0.0329 for complement peers against 0.1233 for substitute peers, so the two groups fall on opposite sides of zero rather than differing in degree. Both measures also separate complement peers from control peers in Table IA.4.[12] This suggests that the releasing firm shares a demand cycle with substitute peers and not with complement peers.

One outside measure should separate the two groups in the opposite direction, and it does so weakly. In the Hoberg and Phillips (2016) vertical data, reported in Table IA.5, complement peers are 1.05 percentage points more likely than substitute peers to be linked to the releasing firm as a supplier or a customer. The difference is not significant ($t = 0.89$). This suggests that the outside

[12]The industry code separates complement peers from control peers by 12.5 percentage points ($t = -5.61$) and sales growth by 0.076 of a correlation ($t = -3.08$), where the TNIC listing separates them by 1.7 percentage points ($t = -2.10$) and the correlation of residual returns not at all. The industry-code gap holds under every combination of fixed effects and clustering, whereas the sales-growth gap does not survive holding the peer firm fixed as well as the release.

data identify the substitute corner as rivals far more sharply than they identify the complement corner as trading partners.

[Insert Table 5 here]

Overall, substitute peers are closer rivals of the releasing firm than complement peers on every outside measure examined, whether the measure comes from product descriptions, an industry code, an accounting statement, or a return series. Table IA.5 repeats the comparison with release fixed effects only and adds six further measures, namely the TNIC similarity score among listed pairs, the correlation of four-factor residual returns, the peer's loading on the releasing firm's residual, and three measures of vertical relatedness. Each separates the two groups in the same direction. Table IA.4 adds control peers. This evidence suggests that the product-market axis separates the exposed groups along the dimension it is built to measure.

3.3 Outcome Variables

I measure the responses of peers in four sets of outcomes: operations, innovation, open-source activity, and default risk and financial policy. The operating outcomes come from Compustat. Sales, capital expenditure, and R&D expenditure are scaled by lagged total assets, and I refer to scaled capital expenditure as investment. Employment per sales is the number of employees per million dollars of sales. The control variables are size, market-to-book, and leverage, where size is the logarithm of total assets. Continuous outcomes are winsorized at the 1st and 99th percentiles. Returns and market capitalization come from CRSP, and factor returns come from the data library of Kenneth French.

Patenting comes from the patent data of Kogan et al. (2017), which link patents to CRSP firms. I count the patents granted to a firm in each fiscal year and, separately, the patents that the USPTO Artificial Intelligence Patent Dataset classifies as artificial intelligence (Giczy et al., 2022). A linked firm-year with no granted patent counts as zero.

Open-source activity comes from the GH Archive event history. GitHub accounts are attributed to firms through a map of 219,916 accounts to 4,467 firms. For each firm-year, I count the events that the firm's accounts record on 1,671 open-source repositories released by large technology firms. The events are forks, issues, issue comments, pull requests, pushes, and watch events. I

exclude events on repositories that the contributing firm owns. Otherwise, the work of a firm's engineers on the firm's own code would count as a response to another firm's release.[13]

Default risk and financial policy enter the tests of alternative explanations. The Merton distance to default solves for the value and volatility of a firm's assets from the market value of its equity, the volatility of its equity, and its debt (Merton, 1974). The O-score default probability is the logistic transformation of the Ohlson (1980) O-score. Leverage is long-term plus current debt over total assets, and cash is cash and short-term investments over total assets. Operating income is operating income before depreciation over lagged total assets, and SG&A is selling, general, and administrative expense over sales.

4 Empirical Results

I examine which peers bear the cost of an open-source release. I estimate the effect at two horizons, because the equity market prices a release within days whereas the sales and spending that the price anticipates take years to appear in the financial statements. I report an announcement return at daily frequency and a stacked difference-in-differences at annual frequency on the same 249 releases. Each estimate compares an exposed group with control peers of the same release, and the difference between the two exposed groups is estimated directly.

4.1 Announcement Returns by Peer Type

I first examine how the equity market prices a release. The announcement return is the first place the two hypotheses can be separated, because a release is public on a known date and the market prices the release before any consequence for revenue can be observed. If a release commoditizes a complement of the releasing firm's core business, complement peers should lose value at the release and substitute peers should hold their value, although the two groups are equally close to the released software. If a release instead raises the value of every firm able to use the software, neither exposed group should lose value, and complement peers should not lose value relative to substitute peers.

[13] A firm with no attributed account records zero events. Engineers who contribute from personal accounts are not counted, which understates the activity of the firms that employ them.

A daily test needs the first day on which the market could trade on the release. Event day 0 is therefore the first trading day after the UTC calendar date on which the repository became public, so the repository is public before the market opens on day 0.[14] A releasing firm publishes many repositories, so the windows around its releases overlap, and averaging abnormal returns across overlapping events treats correlated observations as independent. The announcement return therefore uses calendar-time portfolios (Fama, 1998; Mitchell and Stafford, 2000). For an event window, the complement portfolio holds on each trading day every complement peer of every release whose window contains that day. As a result, a day on which three releases are live counts as one observation. I regress the daily excess return of the portfolio on the Carhart (1997) four factors and multiply the intercept by the number of days in the window. The substitute, control, and releasing-firm portfolios are built in the same way, and the difference between two portfolios is estimated from their daily return spread. Portfolios are weighted by market capitalization at the end of the fiscal year before the release, and standard errors are White heteroskedasticity-robust. Table 6 reports calendar-time portfolio alphas for complement, substitute, control, and releasing-firm portfolios over windows before and after the release.

[Insert Table 6 here]

The complement portfolio earns -0.401 percent on day 0 ($t = -2.97$), -0.625 percent over $[0,+1]$ ($t = -3.34$), and -0.734 percent over $[0,+3]$ ($t = -2.33$).[15] The substitute portfolio in column (2) earns 0.085, 0.058, and 0.042 percent over the same three windows, and none of the three returns differs from zero. The control portfolio in column (3) earns -0.034 percent over $[0,+1]$, and the releasing firm in column (4) does not move either, earning 0.001 percent over $[0,+1]$ ($t = 0.01$). The release day therefore prices the complement portfolio and leaves the other three where they were.

Columns (5) and (6) report the two differences. Over $[0,+1]$, the complement portfolio earns 0.722 percentage points less than the substitute portfolio ($t = -3.55$) and 0.718 percentage points

[14] GitHub records no public event for 112 of the 321 releases, and I date these releases by the creation of the repository. Placing day 0 after the calendar date absorbs part of the resulting imprecision. I chose the rule after comparing six dating rules on the announcement returns. An alternative rule uses the New York trading date and moves releases made public after the 16:00 close to the next trading day. Under the alternative rule, the value-weighted complement return is -0.371 percent over $[0,+3]$ ($t = -1.10$) and is not significant over $[0,+1]$.

[15] With equally weighted portfolios, the complement portfolio earns -0.696 percent over $[0,+1]$ ($t = -3.78$) and 0.685 percentage points less than the substitute portfolio ($t = -3.48$), so the largest complement peers do not drive the result.

less than the releasing-firm portfolio ($t = -2.76$). Neither difference is present before the software becomes public. Over $[-10,-1]$, $[-5,-1]$, and $[-3,-1]$, the complement portfolio earns 0.018, 0.425, and 0.227 percent, and the complement-minus-substitute difference is not significant in any of the three windows. This suggests that a release conveys adverse information about the future sales of complement peers and none about those of substitute peers.[16]

The decline of the complement portfolio is confined to the first four trading days. The complement portfolio earns -0.172 percent over $[0,+5]$ ($t = -0.41$) and -0.400 percent over $[0,+20]$ ($t = -0.45$), and neither longer window is significant.[17]

Table 6 also prices the firm that published the software. The releasing firm gains over longer windows, earning 0.953 percent over $[-10,-1]$ ($t = 1.93$) and 1.544 percent over $[0,+20]$ ($t = 2.44$). The market therefore appears to read a release as a benefit to the firm that publishes it, which is consistent with the gains earlier work finds at firms that engage with, contribute to, and publish open-source software (Emery et al., 2025; Coleman et al., 2025; Conti et al., 2025; Zhao, 2025).[18] Together the four portfolios imply that the market prices a release as a gain for the releasing firm and as a loss for the peers that sell what the released software replaces. The design identifies where the cost falls and not what the releasing firm earns, so the paper makes no claim about the size of that gain.[19]

Every return in Table 6 is cumulated over a window, so the decline over $[0,+3]$ is the same number whether the complement portfolio fell on the release day or three days later. Figure 3 plots the complement, substitute, and releasing-firm portfolios day by day from day -5 to day $+5$. The three portfolios move together before day 0, the complement portfolio separates from the other two on the release day, and it stays below them for the rest of the window. This suggests that the loss is priced on the release day and is not a single-day reaction that reverses at once.

[Insert Figure 3 here]

[16]The event windows cannot be purged of other news about peer firms. Such news would have to occur for complement peers and not for substitute peers, repeatedly across 249 releases, to produce the split between the two groups.

[17]The standard error over $[0,+20]$ is about four times the standard error over $[0,+1]$, so the longer window cannot distinguish a reversal from a decline that persists at its two-day size.

[18]The releasing firms publish many repositories, and a window around one release often contains other releases by the same firm, so the releasing-firm return is not the value of a single release.

[19]Figure IA.1 plots the segment sales of releasing firms around a release. Only thirteen releasing firms report more than one segment, and a release is the firm's own choice, so the figure describes the releasing firms and does not estimate an effect of a release.

Figure 4 plots the difference between the complement portfolio and the releasing-firm portfolio and the difference between the complement portfolio and the substitute portfolio over the same window. Both differences sit near zero on every day before the release and open on day 0. This suggests that the two exposed groups are priced alike until the software becomes public, and that the release is what separates them.

[Insert Figure 4 here]

The windows read so far end within four trading days of the release. A longer horizon asks whether the market gives the loss back once the consequences of a release become clearer. Figure 5 cumulates returns from day -5 to day $+20$. The complement portfolio stays below zero through day 15, apart from days 5 and 10, before turning positive, the substitute portfolio stays near zero throughout, and the releasing-firm portfolio rises over the month. This suggests that the loss to complement peers is priced within the first few trading days, and that the month after the release adds no movement in either peer portfolio this test can detect.

[Insert Figure 5 here]

Overall, the announcement returns separate the two exposed groups on product-market position rather than on technological distance. Complement peers lose value at the release, substitute peers are priced the same before and after it, and the releasing firm gains over the following month. Both exposed groups are equally close to the released software, so the loss belongs to selling what the software replaces and not to being near it. This is the pattern that information transmission between economically linked firms predicts (Lee et al., 2024), with the link here from a release to the firms whose products it displaces. The announcement return is the market's forecast of what a release will do, and the forecast falls entirely on one of the two exposed groups.

4.2 Sales, Investment, and Employment

Next, I examine whether the announcement return corresponds to a contraction in what complement peers sell and spend over the three years after a release. A market reaction is a forecast of future profits, and a return at a firm other than the announcing firm can arise through more than one channel (Lang and Stulz, 1992). If the release displaces the products of complement peers, their

sales should fall, because the released software is free and a fall in the price of one good lowers demand for the goods it replaces (Cournot, 1838; Adner and Lieberman, 2021). A free product has had that effect in other markets, where incumbents responded by changing the prices they charge and the products they sell (Seamans and Zhu, 2014). In turn, investment and employment should fall with sales, because a firm facing a lasting decline in demand for a product contracts the capacity that serves it (Harrigan, 1980). If the release does not displace what complement peers sell, their sales, investment, and employment should not change.

To test these predictions, I estimate a stacked difference-in-differences over event years -3 to $+3$. For peer i of release j in fiscal year t, the specification is:

$$y_{ijt} = \beta_C \left(\text{Post}_{jt} \times \text{Complement}_{ij}\right) + \beta_S \left(\text{Post}_{jt} \times \text{Substitute}_{ij}\right) + \gamma' X_{it} + \alpha_{ij} + \delta_{jt} + \varepsilon_{ijt}, \quad (2)$$

where Post equals one from the fiscal year of the release onward and X_{it} holds the control variables. Each release defines a cohort, and the cohort-by-firm fixed effect α_{ij} absorbs everything constant about a peer within a release, and the cohort-by-year fixed effect δ_{jt} absorbs everything common to the peers of a release in a year. Accordingly, the coefficients compare exposed peers with control peers of the same release in the same fiscal year.

I estimate each outcome in three ways. Column (1) of each panel table keeps complement and control peers and estimates β_C alone. Column (2) keeps substitute and control peers and estimates β_S alone. Column (3) keeps all three groups, estimates both coefficients jointly, and reports their difference with a Wald test. Columns (1) and (2) compare each exposed group with control peers matched to that group, and column (3) compares the two exposed groups within one regression.

The control group holds technological distance fixed but not product-market position.[20] The product-market contrast therefore comes from the substitute group, which is as close to the released software as the complement group and at the opposite end of the product-market axis. Accordingly, a difference between the responses of the two exposed groups is informative about product-market position rather than technological proximity.

Because exposed peers differ from control peers before the release, I reweight control peers by entropy balancing (Hainmueller, 2012). The weights match the control group to the exposed group

[20] A control peer is admitted anywhere up to the 75th percentile of both product-market measures, so control peers cover the product-market positions of complement peers and most of the positions of substitute peers.

on ten pre-release moments, seven levels and three changes from $t-3$ to $t-1$, so that pre-release trends are aligned as well as levels.[21] In columns (1) and (2) the target is the exposed group of the column, and in column (3) the two exposed groups together.[22] Standard errors are clustered by release and fiscal year. In the event-study figures, Post is replaced by indicators for event years $t-3$ to $t+3$ in the column (3) specification, with $t-1$ omitted.

Table 7 examines whether complement peers sell less after a release. The dependent variable in equation (2) is sales scaled by lagged total assets. The coefficient on Post $\times$ Complement is -0.0377 in column (1) ($t=-4.37$), and the coefficient on Post $\times$ Substitute is 0.0075 in column (2) ($t=2.15$). In column (3), the two coefficients are -0.0603 ($t=-7.00$) and 0.0122 ($t=3.68$), and their difference is -0.0725 ($t=-8.32$).[23] This implies that complement peers sell less after a release relative to control peers, whereas substitute peers sell more. In economic terms, the column (1) coefficient is a decline of 4.3 percent of the pre-release complement mean of 0.8725.[24]

[Insert Table 7 here]

Figure 6 shows that the decline begins in the year of the release and reaches -0.0664 in year $t+2$ ($t=-4.23$). Before the release, the complement coefficients are 0.0236 in year $t-3$ ($t=1.77$) and 0.0036 in year $t-2$ ($t=0.26$), so complement peers had slightly higher sales than control peers three years before the release. Figure IA.2 plots the complement-minus-substitute difference over event time for every outcome.

[Insert Figure 6 here]

Table 8 examines whether the investment of complement peers falls with their sales, with capital expenditure scaled by lagged total assets as the dependent variable. The coefficient on

[21] The levels are size, leverage, sales, cost of goods sold, gross profit, cash, and capital expenditure, and the changes are in sales, gross profit, and R&D.

[22] Each exposed group is balanced separately in columns (1) and (2) because complement peers are 1.368 log points smaller than substitute peers, and a joint target concentrates the weights on fewer control peers.

[23] With standard errors clustered by release instead of by release and fiscal year, the t-stat of the column (1) coefficient is -3.36.

[24] Sales scaled by lagged total assets rise when assets fall, so Panels A and B of Table IA.6 measure sales in dollars and in logarithms. Complement peers lose 0.037 log points of dollar sales ($t=-1.53$) and 0.0526 log points of sales in logarithms ($t=-5.90$) in column (3), against 0.008 ($t=0.35$) and -0.0057 ($t=-0.64$) in column (1). Part of the column (1) decline in scaled sales therefore reflects faster growth in assets, and the dollar specification does not separate the two exposed groups on its own. Profits fall in both specifications, and in column (1) of Panels A and B of Table IA.7 gross profit falls by 0.0180 of lagged assets ($t=-3.90$) and operating income by 0.0072 ($t=-2.36$).

Post × Complement is −0.0031 in column (1) ($t = -5.71$), and the coefficient on Post × Substitute is 0.0008 in column (2) ($t = 2.50$). The column (3) difference between the two groups is −0.0026 ($t = -4.53$). This implies that complement peers invest less after a release, whereas substitute peers invest slightly more. In economic terms, the column (1) coefficient is a decline of 10.4 percent of the pre-release mean of 0.0297.[25]

[Insert Table 8 here]

Figure 7 plots the event-study coefficients for investment. The complement coefficients are −0.0003 in year $t-3$ and −0.0001 in year $t-2$, and neither is significant. The complement coefficient is −0.0013 in the year of the release ($t = -1.44$) and falls to −0.0026 by year $t+3$ ($t = -2.47$). The decline in investment follows the release rather than a trend that began before it.

[Insert Figure 7 here]

Table 9 examines whether complement peers employ fewer workers for the sales they make, with the number of employees per million dollars of sales as the dependent variable. The coefficient on Post × Complement is −0.246 in column (1) ($t = -5.13$), and the coefficient on Post × Substitute is 0.057 in column (2) ($t = 5.04$). The column (3) difference is −0.311 ($t = -6.74$). Employment at complement peers falls relative to their sales. In economic terms, the column (1) coefficient is a decline of 5.0 percent of the pre-release mean of 4.945.[26]

[Insert Table 9 here]

Figure 8 plots the event-study coefficients for employment per sales. The complement coefficients are 0.056 in year $t-3$ and 0.040 in year $t-2$, and neither is significant. The complement coefficient falls to −0.197 in the year of the release ($t = -3.07$) and to −0.272 by year $t+3$ ($t = -3.37$). However, the substitute coefficient in year $t-3$ is −0.113 ($t = -5.53$), so the positive substitute coefficient in Table 9 continues a trend that began before the release.

[Insert Figure 8 here]

[25] Total investment, which adds R&D and acquisitions net of asset sales (Richardson, 2006), falls by 0.0196 of lagged assets in column (1) of Panel A of Table IA.8 ($t = -3.86$).

[26] In Panels B and C of Table IA.8, the number of employees at complement peers falls by 0.069 log points in column (3) ($t = -8.34$) and by 0.043 in column (1) ($t = -5.39$).

This evidence suggests that a release reduces the scale of the business of complement peers. Complement peers sell less, invest less, and employ fewer workers per dollar of sales, whereas substitute peers show none of the three declines. Accordingly, the announcement return corresponds to a contraction in the operations of complement peers rather than to a revision of expectations alone.

4.3 R&D, Patenting, and Engagement with Open Source

Next, I examine how complement peers change their own development after a release. A firm whose product is replaced by free software can keep developing the product, or it can reduce its own development and work with open-source code instead.[27] If complement peers keep developing their products, their R&D and patenting should not fall. If complement peers reduce their own development, their R&D and patenting should fall, because free code substitutes for engineering the firm would otherwise fund, and adopting open-source software raises productivity at the firms that use it (Nagle, 2019). In turn, their activity on open-source repositories should rise, because a firm that depends on external code contributes to maintaining it (Conti et al., 2025; Lin and Maruping, 2022). Patents and GitHub contributions are counts, so I estimate those outcomes by Poisson pseudo-maximum likelihood with the same fixed effects (Cohn et al., 2022), and their coefficients are in log points.

Table 10 examines whether complement peers reduce their own development, with R&D expenditure scaled by lagged total assets as the dependent variable. The coefficient on Post $\times$ Complement is -0.0036 in column (1) ($t = -2.41$), and the coefficient on Post $\times$ Substitute is -0.0011 in column (2) ($t = -1.61$). In column (3), the substitute coefficient is -0.0003 ($t = -0.45$), and the difference between the two groups is -0.0047 ($t = -3.01$). This implies that complement peers reduce R&D by more than substitute peers. In economic terms, the column (1) coefficient is a decline of 3.3 percent of the pre-release mean of 0.1114.

[Insert Table 10 here]

Figure 9 plots the event-study coefficients for R&D. The complement coefficients are 0.0011 in year $t-3$ and -0.0006 in year $t-2$, and neither is significant. The complement coefficient is

[27] Open-source work also changes the tasks of a firm's engineers and the projects they contribute to (Heath et al., 2023; Mkrtchyan et al., 2025; Gupta et al., 2024).

-0.0027 in the year of the release ($t = -1.21$) and -0.0081 in year $t+2$ ($t = -2.91$). The decline in R&D begins in the year of the release and is significant from year $t+2$.

[Insert Figure 9 here]

Table 11 examines whether complement peers patent less after a release, with the number of patents granted in the fiscal year as the dependent variable. In Panel A, the coefficient on Post $\times$ Complement is -0.227 in column (1) ($t = -4.15$), the coefficient on Post $\times$ Substitute is 0.198 in column (2) ($t = 9.43$), and the column (3) difference is -0.336 ($t = -6.72$). In Panel B, which counts patents classified as artificial intelligence, the two coefficients are -0.210 ($t = -4.36$) and 0.053 ($t = 2.50$). This implies that complement peers are granted fewer patents after a release, whereas substitute peers are granted more. In economic terms, the complement coefficients are declines of 20 percent in all patents and 19 percent in artificial intelligence patents.[28]

[Insert Table 11 here]

Figure 10 plots the event-study coefficients for both patent counts. In Panel A, the complement coefficients before the release are 0.010 and -0.011, and the complement coefficient falls to -0.253 by year $t+3$ ($t = -2.39$). In Panel B, the complement coefficients before the release are 0.001 and -0.017, and the complement coefficient falls to -0.286 by year $t+3$ ($t = -3.18$). However, the substitute coefficient in year $t-3$ is -0.092 in Panel A ($t = -3.44$), so part of the substitute increase continues a trend that began before the release.

[Insert Figure 10 here]

Table 12 examines whether complement peers raise their activity on open-source repositories, with the number of GitHub events on repositories the firm does not own as the dependent variable. The coefficient on Post $\times$ Complement is 0.556 in column (1) ($t = 2.59$), the coefficient on Post $\times$ Substitute is 0.253 in column (2) ($t = 7.30$), and the column (3) difference is 0.302 ($t = 1.29$). This implies that complement peers increase their activity on open-source repositories,

[28] A patent is typically granted two to three years after the application, so part of the decline in grants reflects applications filed before the release. Dated by application in Table IA.9, patents of complement peers fall by 0.092 log points in column (3) ($t = -2.34$) and by 0.015 in column (1) ($t = -0.32$), and the difference from substitute peers is -0.111 ($t = -2.90$). The artificial-intelligence split does not survive the change of dating: substitute peers fall by 0.096 log points ($t = -5.19$) against 0.068 at complement peers ($t = -1.34$).

as substitute peers do, but the two exposed groups are no longer separated. In economic terms, the column (1) coefficient is an increase of 74 percent over a pre-release mean of 20.194 events a year, against an increase of 29 percent for substitute peers. Active contributions, which exclude watch events, rise by 113 percent at complement peers and by 21 percent at substitute peers.[29]

[Insert Table 12 here]

Figure 11 plots the event-study coefficients for GitHub contributions. The complement coefficients are -0.107 in year $t-3$ ($t=-0.45$) and 0.256 in year $t-2$. The complement coefficient rises to 0.456 in the year of the release ($t=1.46$) and to 0.811 in year $t+1$ ($t=3.71$). However, the substitute coefficients are -0.317 and -0.197 in years $t-3$ and $t-2$ ($t=-5.83$ and -3.24), so the substitute increase partly continues a trend that began before the release.

[Insert Figure 11 here]

This evidence suggests that complement peers respond to a release by shifting engineering effort from their own development toward open-source software. Complement peers reduce R&D and patenting and increase their activity on open-source repositories, and they differ from substitute peers on all three margins. Because the GitHub measure covers the open-source repositories of large technology firms in general, the measure reflects a shift toward open source rather than work on the release that exposed the peer.

4.4 Financial Distress and Cost Cutting

Finally, I examine whether financial distress or cost cutting explains the contraction of complement peers. A firm can sell, invest, and employ less because it is in distress or because it restructures for reasons unrelated to a release. If complement peers are in distress, their default risk should rise. If complement peers cut costs, their operating income should rise relative to assets and their SG&A should fall relative to sales. If complement peers instead lose sales they cannot replace, their operating income should fall and their SG&A should rise relative to sales, because

[29] In Table IA.10, active contributions, which are forks, issues, comments, pull requests, and pushes, rise by 0.757 log points at complement peers ($t=2.42$) and by 0.188 at substitute peers ($t=3.62$), and the column (3) difference is 0.454 ($t=1.39$, Wald $p=0.164$). Watch events rise by 0.249 at complement peers ($t=3.54$) and by 0.314 at substitute peers ($t=14.56$).

overhead is committed ahead of the sales it supports and does not fall with them within three years. The last two predictions are the two sides of one test, because cost cutting and an unreplaceable loss of sales move margins in opposite directions. I estimate both tests with the column (3) specification, which includes all three groups.

Table 13 tests whether the default risk of complement peers rises, with the Merton distance to default and the Ohlson O-score default probability as the dependent variables. The coefficient on Post $\times$ Complement is -0.166 for the Merton distance to default ($t = -1.38$) and -0.0157 for the O-score default probability ($t = -3.17$). The coefficient on Post $\times$ Substitute is -0.247 for the distance to default ($t = -3.96$) and 0.0012 for the O-score default probability ($t = 0.57$), and the difference between the two groups is not significant for the distance to default, with a Wald p-value of 0.494, while for the O-score default probability it is -0.0169 ($t = -3.33$, Wald $p = 0.001$). This implies that default risk does not rise at complement peers, and that on the O-score measure it falls.[30]

[Insert Table 13 here]

Table 14 examines whether complement peers cut costs after a release and how they adjust their financial policy. The dependent variables are leverage, cash, operating income, and SG&A. The coefficients on Post $\times$ Complement are -0.0115 for operating income ($t = -3.72$) and 0.0175 for SG&A ($t = 2.73$), whereas substitute peers raise operating income by 0.0054 ($t = 3.45$) and show no change in SG&A (-0.0034, $t = -1.55$). For financial policy, the coefficients on Post $\times$ Complement are -0.0148 for leverage ($t = -3.64$) and 0.0094 for cash ($t = 2.65$), whereas substitute peers raise leverage by 0.0271 ($t = 16.02$) and reduce cash by 0.0309 ($t = -19.66$). This implies that complement peers reduce debt and add cash while earning less and spending more on overhead relative to sales. In economic terms, operating income falls by 18 percent of its pre-release mean of 0.0641 and SG&A rises by 3.4 percent of its mean of 0.5118, while leverage falls by 6.3 percent of its mean of 0.2348.

[Insert Table 14 here]

[30] Among the seven further measures of distress and restructuring in Table IA.11, complement peers show no change in special items or acquisitions and a marginal decline in debt issuance. The Altman Z-score (Altman, 1968) and the Zmijewski probability (Zmijewski, 1984) indicate lower default risk at complement peers, whereas the failure probability of Campbell et al. (2008) rises by 0.0100 ($t = 1.78$), against 0.0158 for substitute peers.

This evidence suggests that neither financial distress nor cost cutting accounts for the contraction of complement peers. Lower leverage and higher cash are difficult to reconcile with distress, and lower operating income with higher SG&A relative to sales is difficult to reconcile with cost cutting. Accordingly, the contraction appears more consistent with a loss of sales that complement peers cannot offset by cutting costs.

Overall, the results support the hypothesis that open-source releases harm complement peers and not substitute peers. Complement peers lose value at the release, sell less, invest less, and employ fewer workers per dollar of sales over the following three years. Complement peers also reduce R&D and patenting and increase their activity on open-source repositories. Substitute peers neither lose value nor contract, although they are as close to the released software as complement peers. Accordingly, the evidence suggests that a release takes business from the firms that sell the product it makes free. The evidence points to sales that complement peers cannot replace rather than to distress or cost cutting.

5 Robustness and Discussion

I examine whether the results come from exposure to the released software or from the way the classification and the panel are built. The classification is built in this paper, and the panel stacks releases whose peer pools overlap. If the estimates reflect exposure to the released software, placebo labels drawn from the same peer pools should not reproduce them, and reasonable changes to the classification, the control group, and the sample should not remove them. If the estimates are instead mechanical, placebo labels should produce coefficients of similar size.

5.1 Placebo Labels and Alternative Specifications

I first examine whether placebo labels reproduce the panel estimates. Table IA.12 reports a randomization test. In each of 100 draws the complement, substitute, and control labels are reshuffled among the peers of each release, and column (1) is re-estimated on the reshuffled complement and control peers.[31] For sales, the placebo coefficients have a mean of -0.0009 and a standard deviation of 0.0127, against an observed coefficient of -0.0440. For every outcome, the share of

[31] Each draw keeps the number of peers in each group, and the reshuffled regressions are estimated without weights.

draws whose placebo $|t|$ is at least as large as the observed $|t|$ is 0.040 or less. This implies that placebo labels drawn from the same peer pools do not reproduce the estimates.

However, the same draws show that the conventional t-statistics overstate precision. If the standard errors were correct, 5 percent of the draws would have $|t|$ above 1.96, whereas the actual share ranges from 0.100 for investment to 0.300 for patents. The same peer firm enters the pools of many releases, and clustering by release and fiscal year does not capture the resulting correlation. Because they do not rely on the conventional standard errors, the placebo p-values provide the more reliable inference for column (1).

Table IA.13 examines whether the estimates depend on how the classification is built. The two measures behind each axis agree only in part, with mean rank correlations of 0.600 for technology and 0.452 for product market. A classification built on a single technology measure therefore keeps about 40 percent of the complement peers. Panel A moves one constant of the rule at a time, and Panel B one screen at a time.[32] The panel coefficients keep their sign and are significant at the 10 percent level or better in all fifty estimates of Panel A and in twenty-one of the twenty-five estimates of Panel B.[33] Sales, investment, and employment per sales are significant in every row of both panels. The four exceptions are R&D and GitHub contributions, which both lose significance when the technology axis uses the goods-text score alone.[34]

The complement return over $[0,+1]$ is significant in eight of the nine alternatives in Panel A and in all five rows of Panel B. The exception narrows the exposed groups to the top and bottom quarter, where the return is -0.492 percent ($t=-1.43$) on 154 releases. Widening them to the top and bottom 40 percent yields -0.249 percent ($t=-1.86$). This implies that the panel estimates do not depend on the constants of the rule, and that only R&D and GitHub contributions are sensitive to the screens. The announcement return, by contrast, is sensitive to the size of the exposed groups.

Table IA.14 examines whether the estimates depend on how control peers are weighted, by

[32]The constants are the width of the exposed corners, the technology weights, the thresholds that define the control group, the vertical filter, and the presence of the releasing firms in the peer pools. The screens are the peer floor, the number of complement peers a release must have, the number of measures behind the technology axis, the width of the control group, and the composition of the peer pool.

[33]With every releasing firm removed from all peer pools, the complement return over $[0,+1]$ is -0.434 percent ($t=-2.14$) and the sales coefficient is -0.0515 ($t=-5.79$). With the vertical filter removed, so that peers reporting the releasing firm as a customer remain in the pool, they are -0.548 percent ($t=-2.90$) and -0.0607 ($t=-7.01$).

[34]GitHub contributions also lose significance when every peer outside a corner enters the control group, and R&D is 0.0007 ($t=0.54$) when semiconductors are excluded from the peer pool.

replacing the control group of column (1) in six ways.[35] Every coefficient keeps its sign and all but one remain significant, and the sales coefficient ranges from -0.0249 under inverse propensity weights to -0.0616 under balancing on both exposed groups. The exception is R&D under propensity score matching, at -0.0027 ($t = -1.49$), where the effective control sample falls to 8 percent of control peers and the two patent counts cannot be estimated at all. This implies that the estimates do not depend on entropy balancing, apart from R&D and patenting under propensity score matching.

Table IA.15 examines whether the estimates depend on the control sample, the clustering, the fixed effects, or the sample period, by changing one element of column (1) at a time. Dropping every control peer that is an exposed peer of another release in the same years raises the magnitude of the coefficients on sales, investment, R&D, and both patent counts. Employment per sales becomes -0.0384 and loses significance ($t = -0.70$). R&D is the one coefficient that weakens elsewhere, losing significance under clustering by release ($t = -1.55$) and becoming -0.0018 ($t = -1.29$) under industry-by-year fixed effects, where both patent counts and GitHub contributions shrink but remain significant. Dropping the releases of 2020 and 2021, the releases of 2023 and 2024, or all releases after 2019 changes little.

Table IA.16 examines whether one releasing firm produces the results, by dropping each firm with at least five releases in turn. Every row keeps the sign of the complement return over $[0, +1]$ and of every panel coefficient, and each is significant at the 10 percent level or better. Without Microsoft, the largest releasing firm, the complement return is -0.600 percent ($t = -2.59$) and the sales coefficient is -0.0567 ($t = -5.73$).[36] This implies that no single releasing firm produces the results.

This evidence suggests that the panel estimates reflect exposure to the released software rather than the construction of the classification.

[35] The six are entropy balancing on both exposed groups, entropy balancing on levels only, no weights, inverse propensity weights, coarsened exact matching (Iacus et al., 2012), and propensity score matching to three control peers.

[36] On the 70 releases of Microsoft alone the four panel coefficients other than R&D keep their sign and significance, while neither announcement return is significant.

5.2 Alternative Motives for Releasing

Next, I consider other motives for releasing. Standard-setting, recruiting engineers, building an ecosystem, and signaling the quality of the releasing firm are all plausible motives, and nothing in this paper rules them out. However, each motive predicts a different pattern across the two exposed groups. If a release signals the quality of the releasing firm (Alexy et al., 2013), the prices of both exposed groups should move in the same direction, whereas the two groups respond differently. If a release serves recruiting or ecosystem-building, the firms that build on the software should benefit from wider adoption (Lerner and Tirole, 2002; Boudreau, 2010), whereas complement peers lose value. If a release is classical predation, the harm should reach the firms that compete with the releasing firm for the same revenue (Bolton and Scharfstein, 1990), whereas substitute peers do not lose value.

Standard-setting is the motive the design constrains least. A firm that establishes a new standard makes products built to the previous standard less valuable (Farrell and Saloner, 1985), which is close to what complement peers experience, and it would produce the same asymmetry with no intent to displace anyone. The shift of complement peers toward open-source code is equally consistent with standard-setting. This paper therefore does not separate a release intended to commoditize a complement from a release intended to set a standard whose side effect is commoditization. Because the consequences are the same under either motive, the evidence is about consequences and not about intent.

Overall, the robustness results support the interpretation that the estimates reflect exposure to the released software rather than the mechanics of the classification. Placebo labels do not reproduce the estimates, and the panel estimates survive alternative classifications, control groups, specifications, and samples. However, the announcement return is sensitive to the size of the exposed groups and is short-lived, and the conventional standard errors overstate precision. Accordingly, the panel evidence on sales, investment, employment, and innovation is stronger than the announcement return alone.

6 Conclusion

Large technology firms give away finished software as a matter of course. Alphabet published Kubernetes, Meta published PyTorch, and Microsoft published the compiler for TypeScript, and the organizational accounts of 591 publicly traded firms hold 151,085 public GitHub repositories. Recent work measures the gains at the firms that adopt open-source software, at the firms that publish it, and in the economy as a whole. However, the losses have not been measured systematically.

By classifying the peer firms closest to each released repository into the firms that sell what the software replaces and the firms that sell where the release competes, I measure where the cost of a release falls. First, I find that complement peers lose 0.734 percent of market value over the four trading days that begin with a release, while substitute peers lose nothing. Second, over the following three years complement peers sell less relative to their assets, invest less, and employ fewer workers per dollar of sales, and substitute peers change on none of those margins. Third, complement peers cut their own R&D and patenting and more than double their active contributions to open-source repositories. Finally, neither financial distress nor cost cutting accounts for the contraction, because complement peers hold less debt and more cash while their operating income falls and their overhead rises relative to sales. This implies that the cost of a release falls on the firms selling what the release gives away and not on the firms competing with the releasing firm.

This paper contributes to three strands of literature. First, it adds to the empirical literature on corporate open source. That literature has asked what a release is worth to the firm that publishes it and to the firms that adopt it, and has treated publication as an act of value creation. I extend it to the firms on the other side of a release and show that publication also moves value, away from the firms whose product the released software replaces. The unit of analysis is the firm being commoditized.

Second, it contributes to the literature on how the technology of one firm reaches another firm. Work in that literature weighs the knowledge a firm's research gives away against the rivalry the research creates, and concludes that the knowledge dominates on average. Exposure there is located by how similar two firms are, and similarity points here at the group that does not lose. I measure exposure from the released repository instead, which attaches exposure to an event rather than to a pair of firms.

Third, it contributes to the literature on platform owners and the firms that build on their platforms. That work studies an owner that enters a complementor's market, whereas a release takes revenue out of a market without entering it and without changing any price the releasing firm charges. I show which firms that revenue is taken from, and the same design reaches open-weight model releases and free tiers of paid software. Whether conduct of this kind warrants regulatory attention turns on consumer effects this design does not measure.

References

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

Adner, Ron, and Marvin Lieberman, 2021, Disruption through complements, *Strategy Science* 6, 91–109.

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

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

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

Altman, Edward I., 1968, Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, *The Journal of Finance* 23, 589–609.

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

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

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

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

Bolton, Patrick, and David S. Scharfstein, 1990, A theory of predation based on agency problems in financial contracting, *American Economic Review* 80, 93–106.

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

Campbell, John Y., Jens Hilscher, and Jan Szilagyi, 2008, In search of distress risk, *The Journal of Finance* 63, 2899–2939.

Carhart, Mark M, 1997, On persistence in mutual fund performance, *The Journal of Finance* 52, 57–82.

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

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

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

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

Cohn, Jonathan B., Zack Liu, and Malcolm I. Wardlaw, 2022, Count (and count-like) data in finance, *Journal of Financial Economics* 146, 529–551.

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

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

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

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

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

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

Fama, Eugene F., 1998, Market efficiency, long-term returns, and behavioral finance, *Journal of Financial Economics* 49, 283–306.

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

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

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.

Giczy, Alexander V., Nicholas A. Pairolero, and Andrew A. Toole, 2022, Identifying artificial intelligence (AI) invention: A novel AI patent dataset, *The Journal of Technology Transfer* 47, 476–505.

Graham, Stuart J. H., Galen Hancock, Alan C. Marco, and Amanda Fila Myers, 2013, The uspto trademark case files dataset: Descriptions, lessons, and insights, *Journal of Economics & Management Strategy* 22, 669–705.

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

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

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

Hainmueller, Jens, 2012, Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies, *Political Analysis* 20, 25–46.

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

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

Heath, Davidson, and Christopher Mace, 2020, The strategic effects of trademark protection, *The Review of Financial Studies* 33, 1848–1877.

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

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

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

Iacus, Stefano M., Gary King, and Giuseppe Porro, 2012, Causal inference without balance checking: Coarsened exact matching, *Political Analysis* 20, 1–24.

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

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

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

Lang, Larry H. P., and René M. Stulz, 1992, Contagion and competitive intra-industry effects of bankruptcy announcements, *Journal of Financial Economics* 32, 45–60.

Lee, Charles M. C., Terrence Tianshuo Shi, Stephen Teng Sun, and Ran Zhang, 2024, Production complementarity and information transmission across industries, *Journal of Financial Economics* 155, 103812.

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

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

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

Mendonça, Sandro, Tiago Santos Pereira, and Manuel Mira Godinho, 2004, Trademarks as an indicator of innovation and industrial change, *Research Policy* 33, 1385–1404.

Merton, Robert C., 1974, On the pricing of corporate debt: The risk structure of interest rates, *Journal of Finance* 29, 449–470.

Mitchell, Mark L., and Erik Stafford, 2000, Managerial decisions and long-term stock price performance, *Journal of Business* 73, 287–329.

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

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

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

Ohlson, James A., 1980, Financial ratios and the probabilistic prediction of bankruptcy, *Journal of Accounting Research* 18, 109–131.

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

Richardson, Scott, 2006, Over-investment of free cash flow, *Review of Accounting Studies* 11, 159–189.

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

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

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

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

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

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

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

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

Zmijewski, Mark E., 1984, Methodological issues related to the estimation of financial distress prediction models, *Journal of Accounting Research* 22, 59–82.

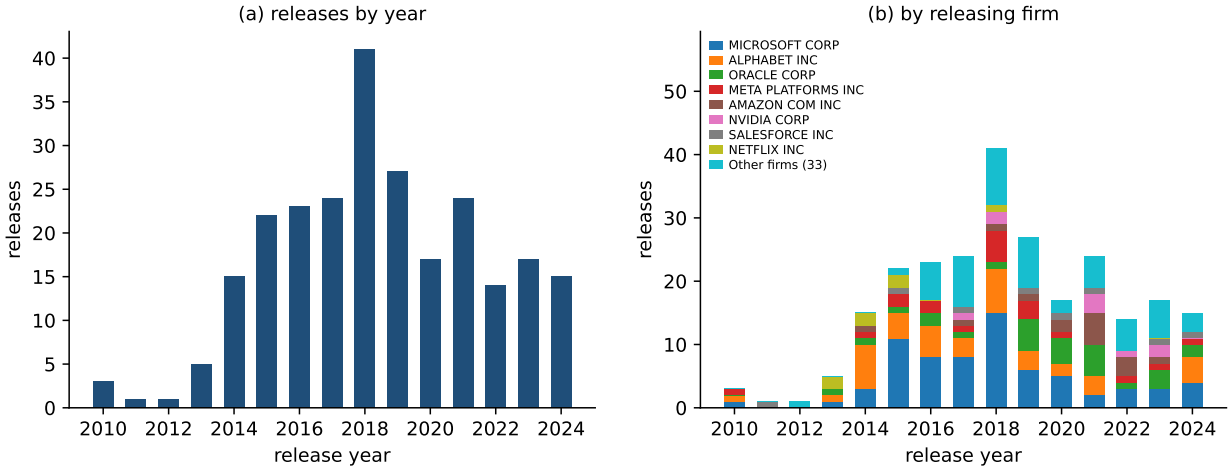

Figure 1. The Estimation Sample

This figure plots the number of releases in the estimation sample by release year, from 2010 to 2024. Panel (a) reports all 249 releases. Panel (b) splits releases by releasing firm; the eight firms with the most releases are shown separately and the remaining firms are pooled.

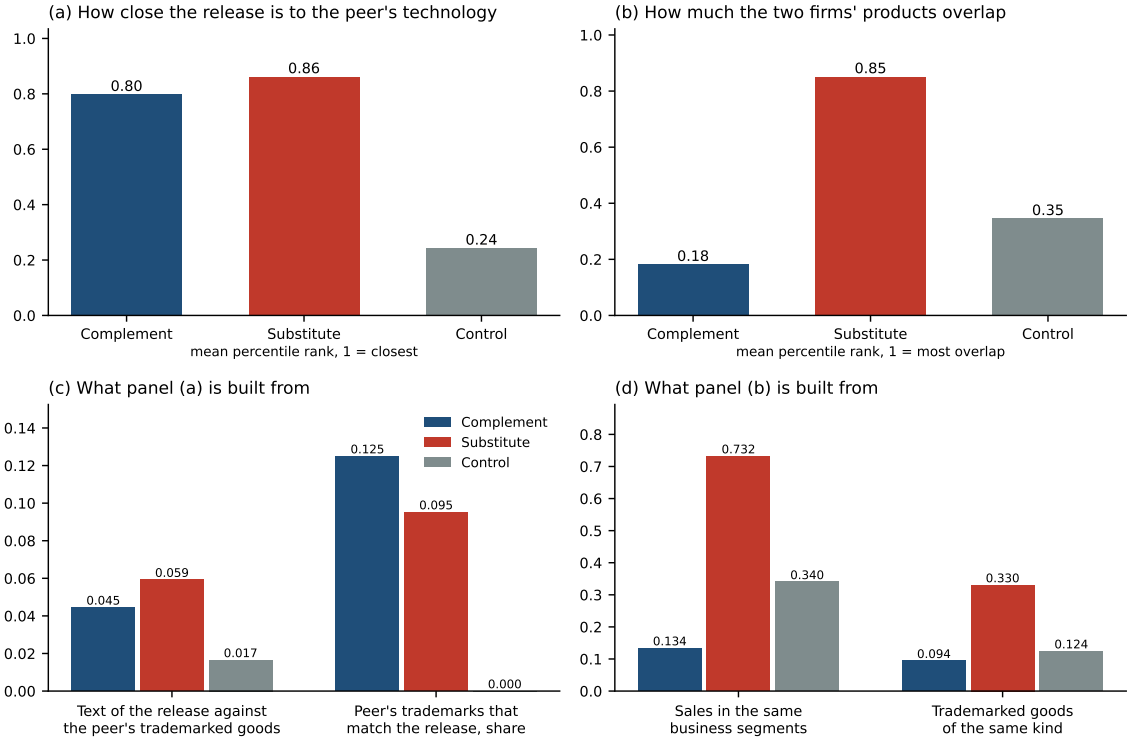

Figure 2. The Three Groups on the Classifier's Own Axes

This figure reports the classification variables by peer group for the 249 releases. Panels (a) and (b) report the mean percentile ranks of complement, substitute and control peers on the technology axis, ranked within the releasing firm's peer pool, and on the product-market axis, ranked within the release; 1 is the closest peer. Panel (c) reports group medians of the two technology measures: the text similarity between the repository and the goods-and-services descriptions of the peer's trademarks, and the share of the peer's live trademarks among the five percent most similar to the repository. Panel (d) reports group medians of the two product-market measures: the segment-sales correlation between the peer and the release's location in segment space, and the trademark goods-and-services similarity between the peer and the releasing firm.

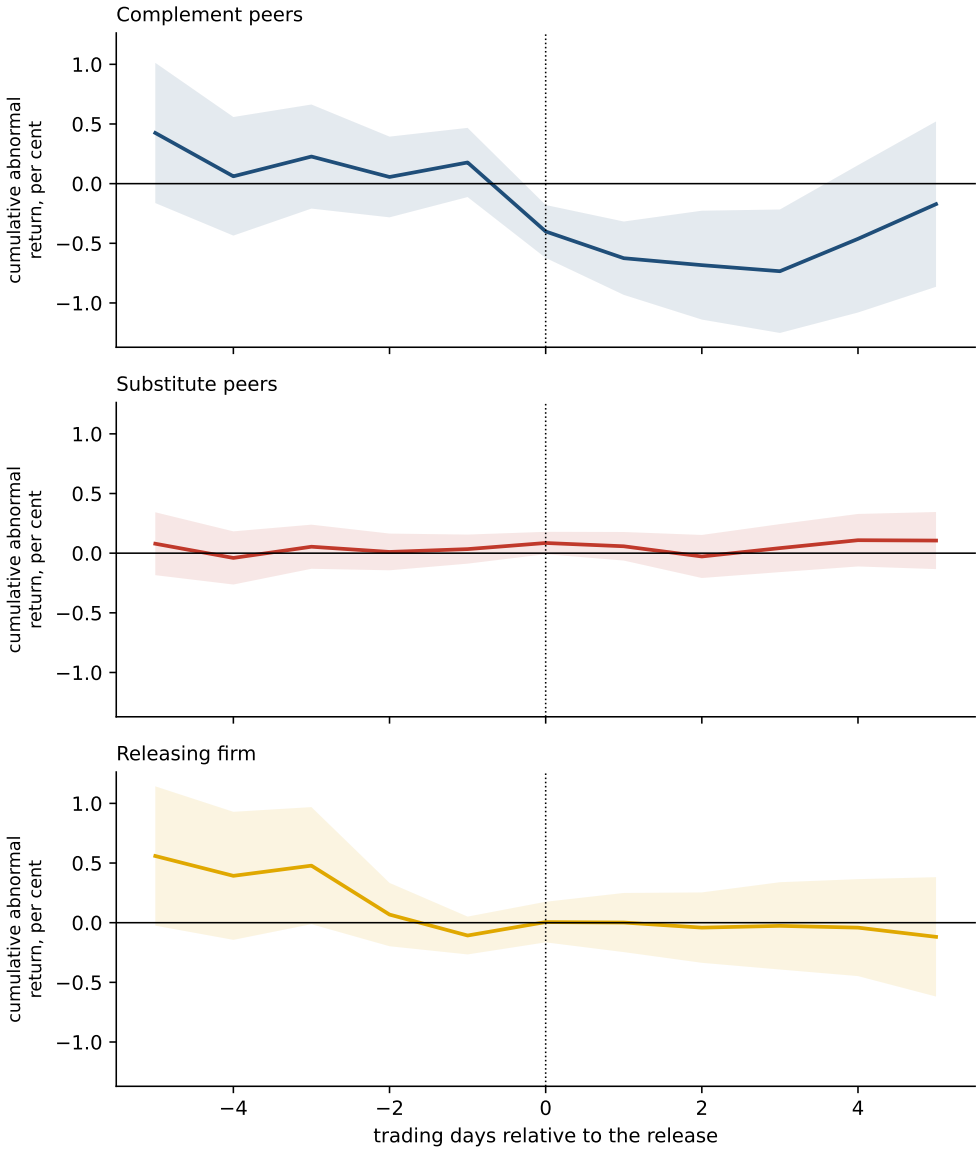

Figure 3. Announcement Returns Day by Day, by Peer Type

This figure plots calendar-time portfolio cumulative abnormal returns for complement peers, substitute peers and releasing firms from event day -5 to $+5$. For a day $d < 0$ the window is $[d, -1]$, and for $d \geq 0$ it is $[0, d]$. Event day 0 is the first trading day after the UTC calendar date on which the repository became public. For each window, the daily excess return of a portfolio holding the firms inside the window is regressed on the Carhart (1997) four factors, and the intercept is multiplied by the number of days in the window. Portfolios are value-weighted by market capitalization. Shaded areas are 90% confidence intervals based on White standard errors.

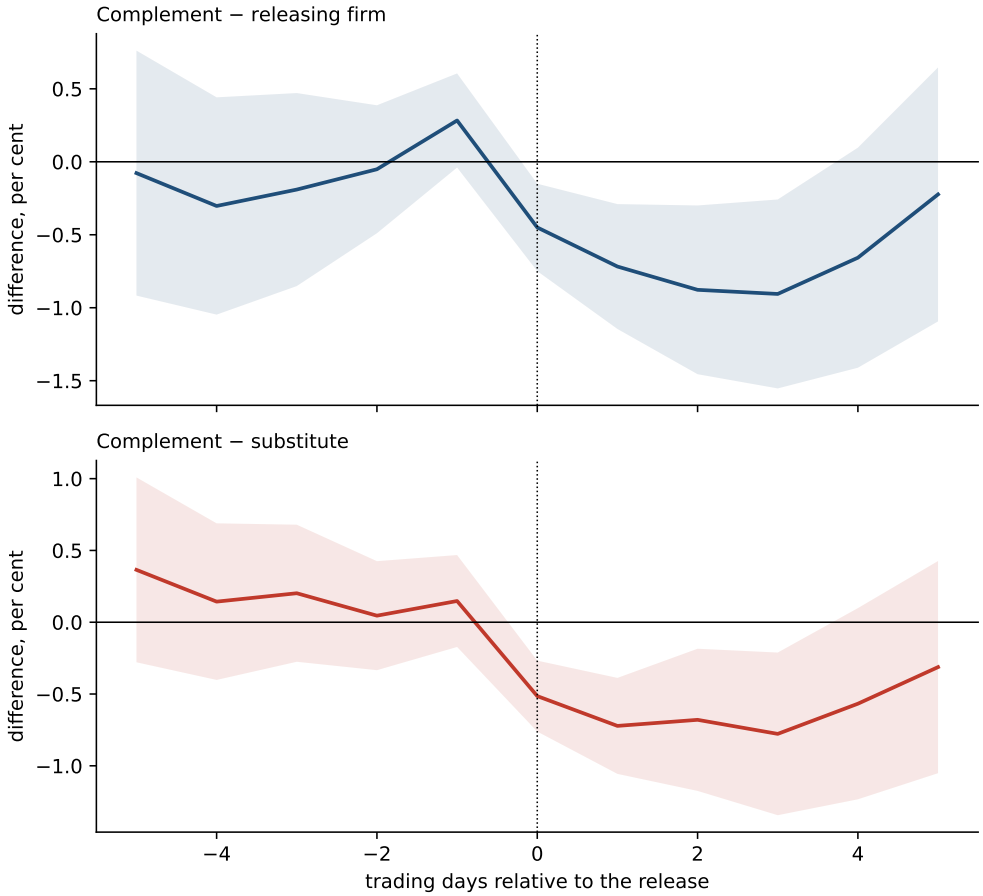

Figure 4. The Complement Portfolio Against the Releasing Firm and Against Substitutes
This figure plots the difference in calendar-time cumulative abnormal returns between the complement portfolio and the releasing-firm portfolio (top) and between the complement portfolio and the substitute portfolio (bottom), from event day -5 to $+5$. For a day $d < 0$ the window is $[d, -1]$, and for $d \geq 0$ it is $[0, d]$. Event day 0 is the first trading day after the UTC calendar date on which the repository became public. For each window, the daily excess return of a portfolio holding the firms inside the window is regressed on the Carhart (1997) four factors, and the intercept is multiplied by the number of days in the window. Portfolios are value-weighted by market capitalization. Shaded areas are 90% confidence intervals based on White standard errors.

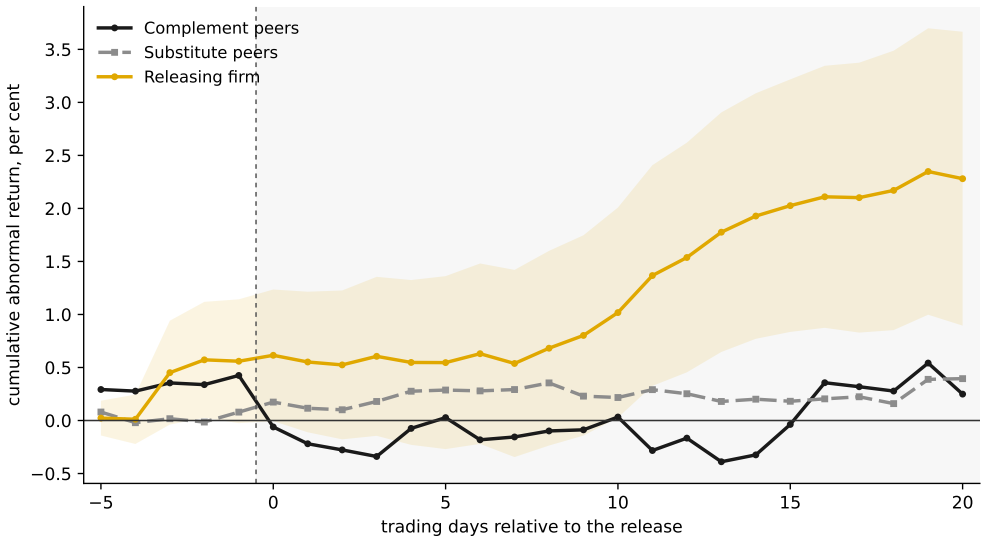

Figure 5. Announcement Returns Day by Day, by Peer Type, Long-Term

This figure plots calendar-time portfolio cumulative abnormal returns for complement peers, substitute peers and releasing firms. Each point is the cumulative abnormal return over the window $[-5, d]$, for d from -5 to $+20$. Event day 0 is the first trading day after the UTC calendar date on which the repository became public. For each window, the daily excess return of a portfolio holding the firms inside the window is regressed on the Carhart (1997) four factors, and the intercept is multiplied by the number of days in the window. Portfolios are value-weighted by market capitalization. The complement and substitute portfolios are drawn in black and grey and are distinguished by line style, and their confidence bands are suppressed because three overlapping bands obscure the paths. The shaded area is a 90% confidence interval for the releasing-firm portfolio based on White standard errors.

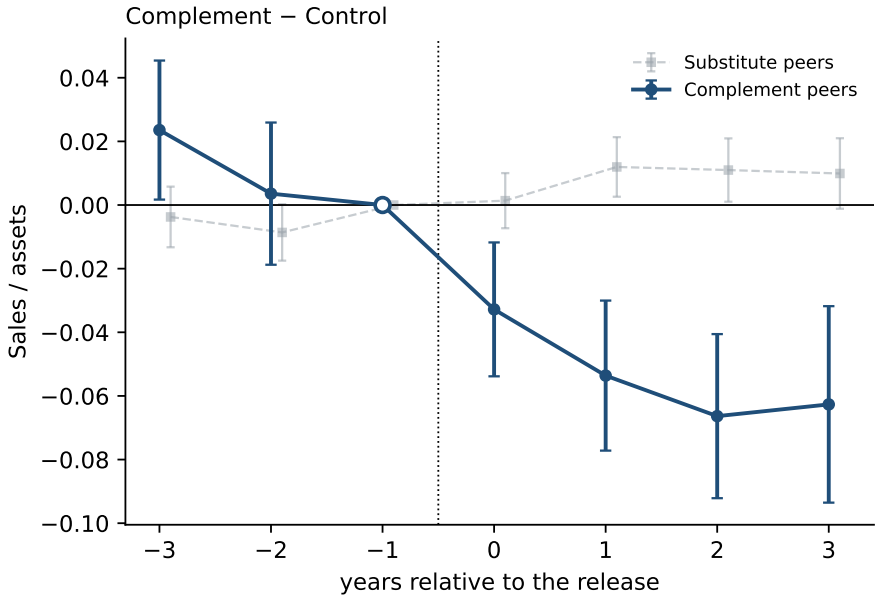

Figure 6. Dynamic Effect on Peer Firms' Sales

This figure plots event-study coefficients for sales scaled by lagged total assets, the outcome of Table 7. Each point is the coefficient on the interaction of an event-year indicator with the complement indicator (circles) or the substitute indicator (squares) from one regression on complement, substitute and control peers, relative to entropy-balanced control peers. Year $t-1$ is the omitted category. Regressions include the controls and fixed effects of Table 7. Bars are 90% confidence intervals based on standard errors clustered at the release-by-fiscal-year level.

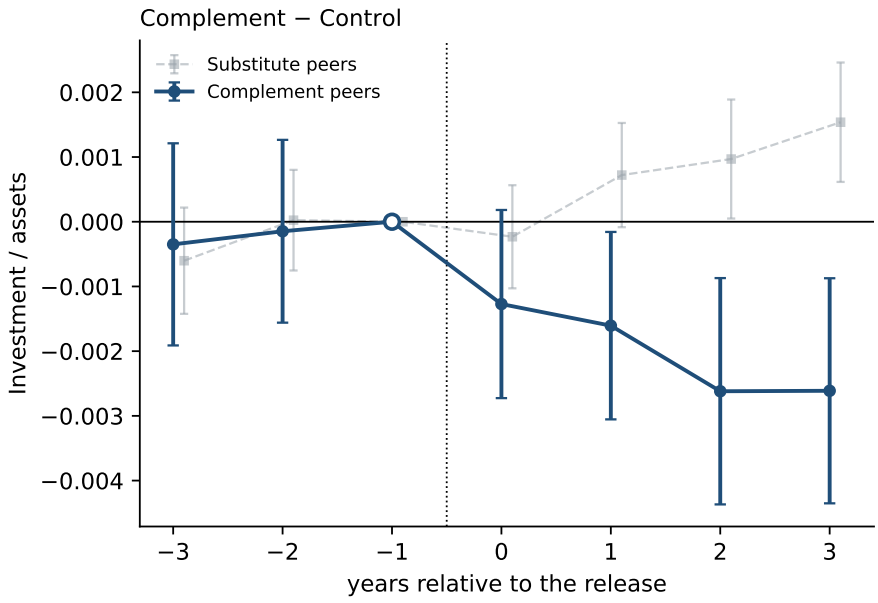

Figure 7. Dynamic Effect on Peer Firms' Investment

This figure plots event-study coefficients for capital expenditure scaled by lagged total assets, the outcome of Table 8. Each point is the coefficient on the interaction of an event-year indicator with the complement indicator (circles) or the substitute indicator (squares) from one regression on complement, substitute and control peers, relative to entropy-balanced control peers. Year $t-1$ is the omitted category. Regressions include the controls and fixed effects of Table 8. Bars are 90% confidence intervals based on standard errors clustered at the release-by-fiscal-year level.

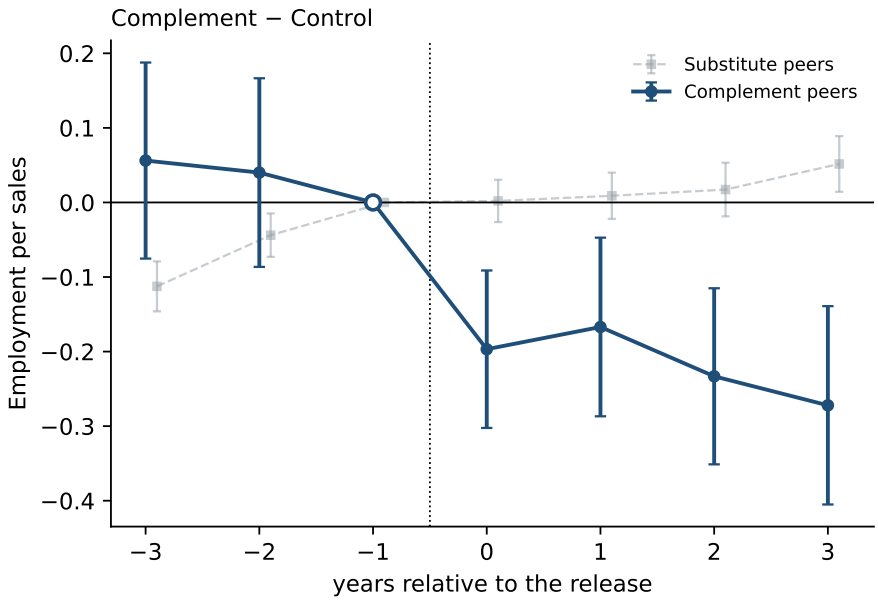

Figure 8. Dynamic Effect on Peer Firms' Employment per Sales

This figure plots event-study coefficients for the number of employees per million dollars of sales, the outcome of Table 9. Each point is the coefficient on the interaction of an event-year indicator with the complement indicator (circles) or the substitute indicator (squares) from one regression on complement, substitute and control peers, relative to entropy-balanced control peers. Year $t-1$ is the omitted category. Regressions include the controls and fixed effects of Table 9. Bars are 90% confidence intervals based on standard errors clustered at the release-by-fiscal-year level.

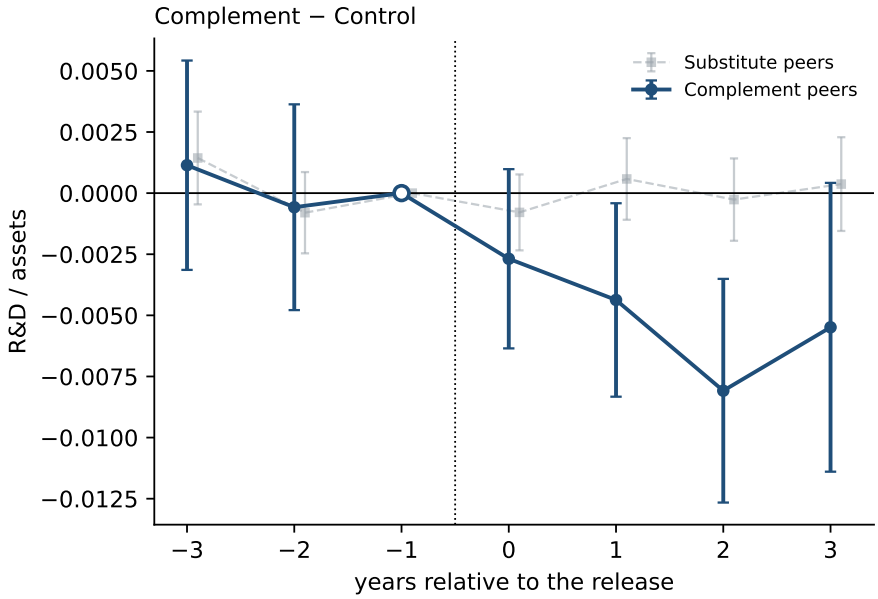

Figure 9. Dynamic Effect on Peer Firms' R&D

This figure plots event-study coefficients for R&D expenditure scaled by lagged total assets, the outcome of Table 10. Each point is the coefficient on the interaction of an event-year indicator with the complement indicator (circles) or the substitute indicator (squares) from one regression on complement, substitute and control peers, relative to entropy-balanced control peers. Year $t-1$ is the omitted category. Regressions include the controls and fixed effects of Table 10. Bars are 90% confidence intervals based on standard errors clustered at the release-by-fiscal-year level.

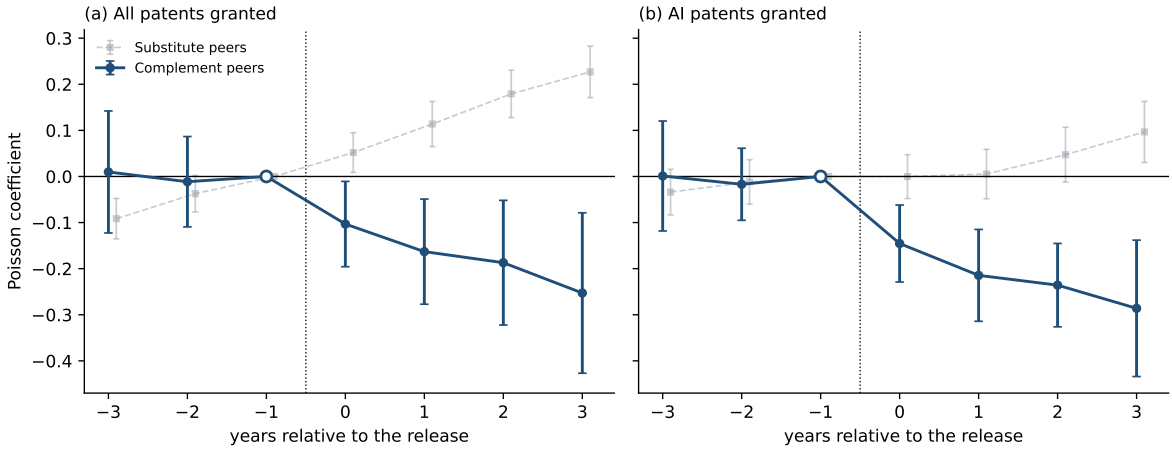

Figure 10. Dynamic Effect on Peer Firms' Patenting

This figure plots event-study coefficients for patents granted, the outcome of Table 11. Each point is the coefficient on the interaction of an event-year indicator with the complement indicator (circles) or the substitute indicator (squares) from one regression on complement, substitute and control peers, relative to entropy-balanced control peers. Panel (a) counts all patents granted and panel (b) patents classified as artificial intelligence. Coefficients are Poisson pseudo-maximum likelihood estimates. Year $t-1$ is the omitted category. Regressions include the controls and fixed effects of Table 11. Bars are 90% confidence intervals based on standard errors clustered at the release-by-fiscal-year level.

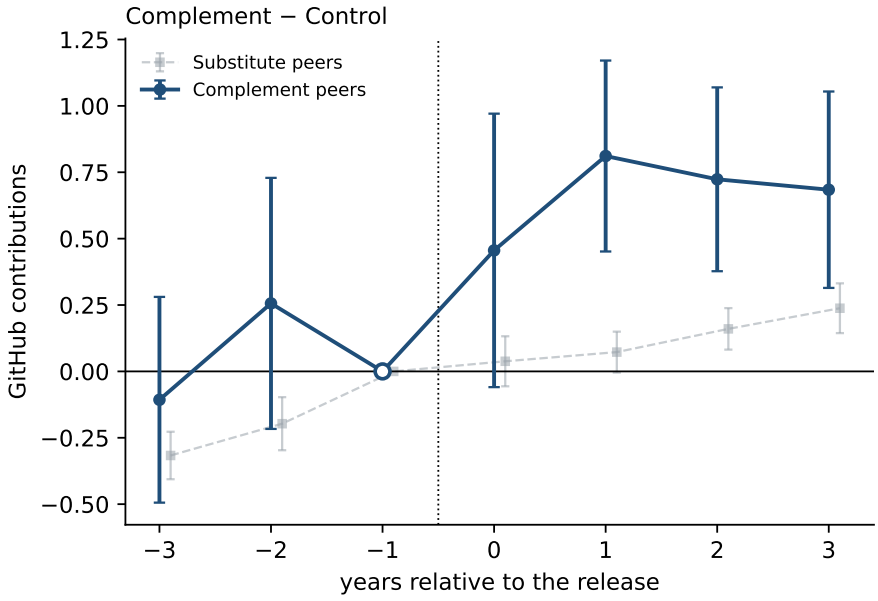

Figure 11. Dynamic Effect on Peer Firms' GitHub Contributions

This figure plots event-study coefficients for GitHub contributions, the outcome of Table 12. Each point is the coefficient on the interaction of an event-year indicator with the complement indicator (circles) or the substitute indicator (squares) from one regression on complement, substitute and control peers, relative to entropy-balanced control peers. Coefficients are Poisson pseudo-maximum likelihood estimates. Year $t-1$ is the omitted category. Regressions include the controls and fixed effects of Table 12. Bars are 90% confidence intervals based on standard errors clustered at the release-by-fiscal-year level.

Table 1. Sample Construction

This table reports the construction of the release sample. Each row applies one screen, in the order applied, to the public GitHub repositories owned by the sample firms. Screens use the GitHub record (release date, fork and template flags, description, detected language and licence), the owner's CRSP share code and listing dates, and the repository's name, description and topics. Rows marked on reading are applied by reading each remaining repository against criteria fixed in advance; a blind second reading of 44 repositories agrees on 43 (Cohen's $\kappa = 0.94$). Each retained repository defines one release, dated by its first public event on GitHub or, when none is recorded, by its creation date. The last row keeps releases with the return and accounting data required for estimation. Removed is the number of repositories dropped by each row, and Remaining is the number left.

	Removed	**Remaining**
Public repositories owned by the sample firms	–	151,085
with a dated release	25,539	125,546
not a fork	21,581	103,965
not a template	782	103,183
with a description	31,311	71,872
with a detected programming language	8,280	63,592
owned by a US common share	8,868	54,724
released while the owner was listed	8,394	46,330
not teaching material or samples, by name	9,282	37,048
not documentation, by name or language	1,846	35,202
not a mirror, configuration or personal repository, by name	2,399	32,803
not a client for the firm's own service	2,698	30,105
with a licence	2,786	27,319
with a software licence, not a content licence	1,285	26,034
with a description of 20 characters or more	520	25,514
not a notebook or shell script	2,596	22,918
not samples, documentation or scratch work, by name or topic	364	22,554
not a duplicate of another repository	289	22,265
not teaching material, an asset dump or a mirror, on reading	7,425	14,840
not a library, plugin, extension or component, on reading	14,519	321
Release events, one per repository kept above	–	321
Estimation sample: releases with the data the estimates require	72	249

Table 2. Well-Known Releases in the Sample

This table lists selected repositories from the 249 releases in the estimation sample. The selection is illustrative and is not used in estimation; Table IA.1 lists all releases. Public is the date on which the repository became public. Rows are sorted by that date.

Repository	Releasing firm	Public
facebook/hhvm	Meta Platforms Inc	2010-01-04
ansible/ansible	International Business Machines	2012-03-07
kubernetes/kubernetes	Alphabet Inc	2014-06-09
bazelbuild/bazel	Alphabet Inc	2014-06-13
microsoft/TypeScript	Microsoft Corp	2014-06-18
mysql/mysql-server	Oracle Corp	2014-09-29
angular/angular	Alphabet Inc	2014-09-30
golang/go	Alphabet Inc	2014-12-09
grpc/grpc	Alphabet Inc	2015-02-27
spinnaker/spinnaker	Netflix Inc	2015-11-17
oracle/graal	Oracle Corp	2016-01-15
microsoft/WSL	Microsoft Corp	2016-04-07
PowerShell/PowerShell	Microsoft Corp	2016-08-19
pytorch/pytorch	Meta Platforms Inc	2017-01-19
microsoft/terminal	Microsoft Corp	2017-08-14
Netflix/iceberg	Netflix Inc	2018-01-05
openjdk/jdk	Oracle Corp	2018-09-18
openvinotoolkit/openvino	Intel Corp	2018-10-16
jax-ml/jax	Alphabet Inc	2018-10-26
corretto/corretto-8	Amazon.com Inc	2018-11-15
microsoft/onnxruntime	Microsoft Corp	2018-12-04
microsoft/PowerToys	Microsoft Corp	2019-05-07
opensearch-project/OpenSearch	Amazon.com Inc	2021-04-13
google-deepmind/alphafold	Alphabet Inc	2021-06-18
google-deepmind/mujoco	Alphabet Inc	2021-08-30
meta-llama/llama	Meta Platforms Inc	2023-02-15
NVIDIA/TensorRT-LLM	Nvidia Corp	2023-10-20
google-deepmind/gemma	Alphabet Inc	2024-02-21

Table 3. Descriptive Summary of the Estimation Sample

This table reports summary statistics for the estimation sample of 249 releases by 41 releasing firms from 2010 to 2024. Panel A describes the released repositories, with one observation per release. Stars, forks, repository size and events on the repository are measured at data collection. Days built privately is the number of days between the repository's creation and its public release. Peers matched is the number of firms in the release's peer pool. Panel B describes the releasing firms, with one observation per firm and fiscal year of release, over the 40 releasing firms with Compustat data in the fiscal year of their release and 119 firm-years. Assets, sales, R&D and market value of equity are in billions of dollars. Cash is cash and short-term investments over total assets, and leverage is long-term plus current debt over total assets.

Panel A: The Released Repository

	Releases	Mean	S.D.	p25	Median	p75
Stars	249	10,083	21,608	495	2,542	9,668
Forks	249	1,877	4,870	115	379	1,271
Repository size (KB)	249	354,223	750,911	17,500	120,196	386,882
Events on the repository	249	76,072	276,164	3,645	13,265	49,837
Days built privately before release	249	304	591	1	48	341
Peers matched to the release	249	323	87	279	319	372

Panel B: The Releasing Firm

	Firm-years	Mean	S.D.	p25	Median	p75
Assets ($bn)	119	126.4	142.5	9.8	99.8	165.1
Sales ($bn)	119	77.7	108.2	6.7	39.8	87.7
R&D ($bn)	119	10.7	14.7	1.1	6.0	13.1
Market value of equity ($bn)	119	455.5	663.9	47.5	195.1	533.0
Cash	119	0.318	0.167	0.181	0.285	0.466
Leverage	119	0.240	0.170	0.115	0.224	0.326

Table 4. The Release-Level Classification of Peer Firms

This table reports the classification of peer firms. The technology axis combines, with weights of 0.67 and 0.33, the percentile ranks within the releasing firm's peer pool of (i) the text similarity between the repository and the goods-and-services text of the peer's trademarks and (ii) the share of the peer's live trademarks among the five percent most similar to the repository. The product-market axis combines the percentile ranks within the release of (i) the Bloom et al. (2013) correlation between the segment-sales shares of the peer and of the fifteen non-releasing firms whose 10-K descriptions are most similar to the repository and (ii) the trademark goods-and-services similarity between the peer and the releasing firm. Complement (substitute) peers are in the top third of the technology axis and the bottom (top) third of both product-market measures; complement peers must in addition share an input-output flow with the releasing firm's industry, in either direction, in the Bureau of Economic Analysis use table. Control peers are at or below the median of the technology axis and the 75th percentile of both product-market measures. Peers that report the releasing firm as a customer are excluded, and releases require at least nine peers. Panel A reports mean percentile ranks, with standard deviations in parentheses. Panel B reports pre-release means over fiscal years $t-3$ to $t-1$ of (release, peer) pairs, with sales, R&D and capital expenditure scaled by lagged total assets. Control means use entropy-balancing weights (Hainmueller, 2012) on the pre-release levels of size, leverage, sales, cost of goods sold, gross profit, cash and capital expenditure and the $t-3$ to $t-1$ changes in sales, gross profit and R&D. Differences are from regressions on a group indicator, with standard errors clustered at the peer-firm level. t-statistics are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Panel A: How the Rule Separates the Peers

	Technology, against the released software	Product market, against the releasing firm
Complement	0.80 (0.09)	0.18 (0.06)
Substitute	0.86 (0.09)	0.85 (0.07)

Panel B: The Three Groups Before the Release

	Group mean			Difference	
	Complement	Substitute	Control	Comp. − ctrl.	Comp. − sub.
Size (log assets)	6.392	7.760	7.819	-1.252***	-1.368***
				(-3.19)	(-4.27)
Market-to-book	4.876	-2.519	4.990	-0.012	7.395
				(-0.00)	(0.82)
Leverage	0.217	0.224	0.232	0.003	-0.006
				(0.09)	(-0.16)
Sales	0.882	0.674	0.696	0.177**	0.208***
				(2.55)	(3.32)
R&D	0.111	0.125	0.103	0.008	-0.014
				(0.68)	(-1.05)
Investment	0.030	0.032	0.030	-0.001	-0.002
				(-0.15)	(-0.48)
(release, peer) observations	890	5,065	14,388		
Distinct peer firms	218	277	414		

Table 5. Complement and Substitute Peers on External Measures

This table compares complement and substitute peers on external measures that are not used in the classification. TNIC rival equals 100 if the peer and the releasing firm are product-market competitors in the Hoberg and Phillips (2016) TNIC data and 0 otherwise. Same SIC4 equals 100 if the two firms report the same four-digit SIC code and 0 otherwise. Sales growth co-movement is the correlation of the two firms' annual sales growth over the six fiscal years ending the year before the release, computed where at least four years overlap. Each column reports an OLS regression of the dependent variable on Complement, an indicator equal to one if the peer is a complement peer and zero if it is a substitute peer, estimated over (release, peer) pairs of complement and substitute peers; control peers are excluded. Mean of dep. var. is the sample mean of the dependent variable for complement peers and for substitute peers. Regressions include release and peer-firm fixed effects. Standard errors are clustered at the peer-firm level. t-statistics are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Dependent variable =	TNIC rival (%)	Same SIC4 (%)	Sales growth co-movement
	(1)	(2)	(3)
Complement	-17.76***	-37.17***	-0.2032**
	(-5.34)	(-5.76)	(-2.35)
Release FE	Yes	Yes	Yes
Peer-firm FE	Yes	Yes	Yes
Observations	6,457	6,457	4,277
Peer firms	425	425	295
Within R^2	0.006	0.023	0.003
Mean of dep. var., complement peers	5.39	15.67	-0.0329
Mean of dep. var., substitute peers	26.49	35.00	0.1233

Table 6. Announcement Returns by Peer Type

This table reports calendar-time portfolio abnormal returns around open-source releases (Fama, 1998; Mitchell and Stafford, 2000). Columns (1) to (4) report portfolios of complement peers, substitute peers, control peers and releasing firms. Columns (5) and (6) report the complement portfolio minus the releasing-firm portfolio and minus the substitute portfolio. Event day 0 is the first trading day after the UTC calendar date on which the repository became public. For each window, the daily excess return of a portfolio holding the firms inside the window is regressed on the Carhart (1997) four factors, and the intercept is multiplied by the number of days in the window. Portfolios are value-weighted by market capitalization. Returns are in percent. Standard errors are White heteroskedasticity-robust. Peers / releases reports the number of peers and releases in each portfolio. The sample covers the 249 releases from 2010 to 2024. t-statistics are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	Portfolio				**Difference**	
Window	(1) Complement	(2) Substitute	(3) Control	(4) Releasing firm	(5) vs. releasing firm	(6) vs. substitute
Before the release						
$[-10,-1]$	0.018	0.069	-0.358*	0.953*	-1.044	-0.038
	(0.03)	(0.31)	(-1.75)	(1.93)	(-1.41)	(-0.07)
$[-5,-1]$	0.425	0.079	0.023	0.559	-0.078	0.365
	(1.19)	(0.49)	(0.18)	(1.57)	(-0.15)	(0.93)
$[-3,-1]$	0.227	0.054	0.042	0.478	-0.190	0.202
	(0.86)	(0.48)	(0.44)	(1.60)	(-0.47)	(0.70)
The release day						
$[0,0]$	-0.401***	0.085	-0.038	0.005	-0.450**	-0.514***
	(-2.97)	(1.50)	(-0.65)	(0.05)	(-2.47)	(-3.43)
After the release						
$[0,+1]$	-0.625***	0.058	-0.034	0.001	-0.718***	-0.722***
	(-3.34)	(0.79)	(-0.43)	(0.01)	(-2.76)	(-3.55)
$[0,+3]$	-0.734**	0.042	-0.115	-0.027	-0.906**	-0.778**
	(-2.33)	(0.35)	(-0.98)	(-0.12)	(-2.30)	(-2.26)
$[0,+5]$	-0.172	0.106	-0.063	-0.119	-0.223	-0.313
	(-0.41)	(0.73)	(-0.40)	(-0.39)	(-0.42)	(-0.70)
$[0,+10]$	0.045	0.066	-0.126	0.446	-0.324	-0.129
	(0.07)	(0.31)	(-0.53)	(1.17)	(-0.41)	(-0.18)
$[0,+20]$	-0.400	0.161	0.215	1.544**	-2.227**	-0.711
	(-0.45)	(0.49)	(0.60)	(2.44)	(-2.03)	(-0.75)
Peers / releases	968 / 249	5435 / 248	25469 / 249	247 / 247		

Table 7. The Effect of an Open-Source Release on Peer Firms' Sales

This table reports stacked difference-in-differences estimates of the effect of open-source releases on peer firms' sales. The dependent variable is sales scaled by lagged total assets. Post equals one from the fiscal year of the release onward. Column (1) compares complement peers and column (2) substitute peers with control peers, each excluding the other exposed group and weighting control peers by entropy balancing on that group. Column (3) includes all three groups, weights control peers on complement and substitute peers together and estimates both coefficients jointly; Complement − substitute is their difference, with a Wald p-value. Weights match the pre-release moments in Table 4. Mean of dep. var. is the pre-release mean of the treated group of each column. Control variables are size, market-to-book and leverage. All regressions include cohort-by-firm and cohort-by-year fixed effects, where each release defines a cohort. Standard errors are clustered at the release-by-fiscal-year level. The sample covers the 249 releases from 2010 to 2024 and, for each release, fiscal years $t-3$ to $t+3$. t-statistics are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Dependent variable =	Sales		
	(1)	(2)	(3)
Post × Complement	-0.0377***		-0.0603***
	(-4.37)		(-7.00)
Post × Substitute		0.0075**	0.0122***
		(2.15)	(3.68)
Size	-0.0855***	-0.0777***	-0.0867***
	(-18.86)	(-16.25)	(-19.04)
Market-to-book	-0.0000	-0.0000*	-0.0000**
	(-1.53)	(-1.66)	(-2.00)
Leverage	-0.1513***	-0.0739***	-0.0809***
	(-13.39)	(-6.45)	(-7.81)
Complement − substitute			-0.0725***
			(-8.32)
Wald p			0.000
Cohort × Firm FE	Yes	Yes	Yes
Cohort × Year FE	Yes	Yes	Yes
Mean of dep. var.	0.8725	0.6668	0.8725
Observations	91,631	112,917	116,949
Adjusted R^2	0.743	0.777	0.757

Table 8. The Effect of an Open-Source Release on Peer Firms' Investment

This table reports stacked difference-in-differences estimates of the effect of open-source releases on peer firms' investment. The dependent variable is capital expenditure scaled by lagged total assets. Post equals one from the fiscal year of the release onward. Column (1) compares complement peers and column (2) substitute peers with control peers, each excluding the other exposed group and weighting control peers by entropy balancing on that group. Column (3) includes all three groups, weights control peers on complement and substitute peers together and estimates both coefficients jointly; Complement − substitute is their difference, with a Wald p-value. Weights match the pre-release moments in Table 4. Mean of dep. var. is the pre-release mean of the treated group of each column. Control variables are size, market-to-book and leverage. All regressions include cohort-by-firm and cohort-by-year fixed effects, where each release defines a cohort. Standard errors are clustered at the release-by-fiscal-year level. The sample covers the 249 releases from 2010 to 2024 and, for each release, fiscal years $t-3$ to $t+3$. t-statistics are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Dependent variable =	Investment		
	(1)	(2)	(3)
Post × Complement	-0.0031***		-0.0017***
	(-5.71)		(-3.10)
Post × Substitute		0.0008**	0.0008***
		(2.50)	(2.81)
Size	0.0032***	0.0002	0.0009***
	(10.28)	(0.75)	(3.59)
Market-to-book	-0.0000***	0.0000**	0.0000**
	(-3.51)	(2.29)	(2.22)
Leverage	-0.0092***	-0.0030***	-0.0041***
	(-15.30)	(-3.91)	(-6.16)
Complement − substitute			-0.0026***
			(-4.53)
Wald p			0.000
Cohort × Firm FE	Yes	Yes	Yes
Cohort × Year FE	Yes	Yes	Yes
Mean of dep. var.	0.0297	0.0305	0.0297
Observations	91,631	112,917	116,949
Adjusted R^2	0.644	0.776	0.755

Table 9. The Effect of an Open-Source Release on Peer Firms' Employment per Sales

This table reports stacked difference-in-differences estimates of the effect of open-source releases on peer firms' employment. The dependent variable is the number of employees per million dollars of sales. Post equals one from the fiscal year of the release onward. Column (1) compares complement peers and column (2) substitute peers with control peers, each excluding the other exposed group and weighting control peers by entropy balancing on that group. Column (3) includes all three groups, weights control peers on complement and substitute peers together and estimates both coefficients jointly; Complement − substitute is their difference, with a Wald p-value. Weights match the pre-release moments in Table 4. Mean of dep. var. is the pre-release mean of the treated group of each column. Control variables are size, market-to-book and leverage. All regressions include cohort-by-firm and cohort-by-year fixed effects, where each release defines a cohort. Standard errors are clustered at the release-by-fiscal-year level. The sample covers the 249 releases from 2010 to 2024 and, for each release, fiscal years $t-3$ to $t+3$. t-statistics are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Dependent variable =	Employment per sales		
	(1)	(2)	(3)
Post × Complement	-0.246***		-0.242***
	(-5.13)		(-5.37)
Post × Substitute		0.057***	0.069***
		(5.04)	(6.06)
Size	-0.191***	-0.027**	-0.019
	(-8.55)	(-2.33)	(-1.62)
Market-to-book	0.000	0.000***	0.000***
	(0.65)	(2.62)	(2.69)
Leverage	0.274***	-0.166***	-0.170***
	(5.07)	(-6.05)	(-6.35)
Complement − substitute			-0.311***
			(-6.74)
Wald p			0.000
Cohort × Firm FE	Yes	Yes	Yes
Cohort × Year FE	Yes	Yes	Yes
Mean of dep. var.	4.945	3.556	4.945
Observations	91,229	112,512	116,534
Adjusted R^2	0.905	0.912	0.915

Table 10. The Effect of an Open-Source Release on Peer Firms' R&D

This table reports stacked difference-in-differences estimates of the effect of open-source releases on peer firms' R&D. The dependent variable is R&D expenditure scaled by lagged total assets. Post equals one from the fiscal year of the release onward. Column (1) compares complement peers and column (2) substitute peers with control peers, each excluding the other exposed group and weighting control peers by entropy balancing on that group. Column (3) includes all three groups, weights control peers on complement and substitute peers together and estimates both coefficients jointly; Complement − substitute is their difference, with a Wald p-value. Weights match the pre-release moments in Table 4. Mean of dep. var. is the pre-release mean of the treated group of each column. Control variables are size, market-to-book and leverage. All regressions include cohort-by-firm and cohort-by-year fixed effects, where each release defines a cohort. Standard errors are clustered at the release-by-fiscal-year level. The sample covers the 249 releases from 2010 to 2024 and, for each release, fiscal years $t-3$ to $t+3$. t-statistics are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Dependent variable =	R&D		
	(1)	(2)	(3)
Post × Complement	-0.0036**		-0.0050***
	(-2.41)		(-3.35)
Post × Substitute		-0.0011	-0.0003
		(-1.61)	(-0.45)
Size	-0.0132***	-0.0167***	-0.0164***
	(-21.84)	(-17.93)	(-19.94)
Market-to-book	0.0000	0.0000	0.0000
	(0.57)	(1.57)	(0.73)
Leverage	0.0133***	-0.0071***	-0.0057***
	(9.13)	(-3.90)	(-3.38)
Complement − substitute			-0.0047***
			(-3.01)
Wald p			0.003
Cohort × Firm FE	Yes	Yes	Yes
Cohort × Year FE	Yes	Yes	Yes
Mean of dep. var.	0.1114	0.1220	0.1114
Observations	91,631	112,917	116,949
Adjusted R^2	0.834	0.837	0.834

Table 11. The Effect of an Open-Source Release on Peer Firms' Patenting

This table reports stacked difference-in-differences estimates of the effect of open-source releases on peer firms' patenting. The dependent variable is the number of patents granted in the fiscal year, all patents in Panel A and those classified as artificial intelligence by Giczy et al. (2022) in Panel B. Coefficients are Poisson pseudo-maximum likelihood estimates. Post equals one from the fiscal year of the release onward. Column (1) compares complement peers and column (2) substitute peers with control peers, each excluding the other exposed group and weighting control peers by entropy balancing on that group. Column (3) includes all three groups, weights control peers on complement and substitute peers together and estimates both coefficients jointly; Complement − substitute is their difference, with a Wald p-value. Weights match the pre-release moments in Table 4. Mean of dep. var. is the pre-release mean of the treated group of each column. Control variables are size, market-to-book and leverage. All regressions include cohort-by-firm and cohort-by-year fixed effects, where each release defines a cohort. Standard errors are clustered at the release-by-fiscal-year level. The sample covers the 249 releases from 2010 to 2024 and, for each release, fiscal years $t-3$ to $t+3$. t-statistics are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)
Panel A: Patents granted			
Post × Complement	-0.227***		-0.166***
	(-4.15)		(-3.23)
Post × Substitute		0.198***	0.170***
		(9.43)	(8.81)
Complement − substitute			-0.336***
			(-6.72)
Wald p			0.000
Mean of dep. var.	32.295	451.758	32.295
Observations	89,951	110,844	114,815
Pseudo R^2	0.010	0.076	0.069
Panel B: AI patents granted			
Post × Complement	-0.210***		-0.209***
	(-4.36)		(-4.83)
Post × Substitute		0.053**	0.045**
		(2.50)	(2.24)
Complement − substitute			-0.253***
			(-6.21)
Wald p			0.000
Mean of dep. var.	23.157	580.205	23.157
Observations	89,951	110,844	114,815
Pseudo R^2	0.035	0.069	0.082
Controls	Yes	Yes	Yes
Cohort × Firm FE	Yes	Yes	Yes
Cohort × Year FE	Yes	Yes	Yes

Table 12. The Effect of an Open-Source Release on Peer Firms' GitHub Contributions

This table reports stacked difference-in-differences estimates of the effect of open-source releases on peer firms' GitHub contributions. The dependent variable is the number of GitHub events (forks, issues, issue comments, pull requests, pushes and watch events) by the firm's accounts on repositories the firm does not own. Coefficients are Poisson pseudo-maximum likelihood estimates. Post equals one from the fiscal year of the release onward. Column (1) compares complement peers and column (2) substitute peers with control peers, each excluding the other exposed group and weighting control peers by entropy balancing on that group. Column (3) includes all three groups, weights control peers on complement and substitute peers together and estimates both coefficients jointly; Complement − substitute is their difference, with a Wald p-value. Weights match the pre-release moments in Table 4. Mean of dep. var. is the pre-release mean of the treated group of each column. Control variables are size, market-to-book and leverage. All regressions include cohort-by-firm and cohort-by-year fixed effects, where each release defines a cohort. Standard errors are clustered at the release-by-fiscal-year level. The sample covers the 249 releases from 2010 to 2024 and, for each release, fiscal years $t-3$ to $t+3$. t-statistics are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Dependent variable =	GitHub contributions		
	(1)	(2)	(3)
Post × Complement	0.556***		0.549**
	(2.59)		(2.29)
Post × Substitute		0.253***	0.247***
		(7.30)	(7.33)
Size	0.419***	-0.051*	0.025
	(9.99)	(-1.87)	(0.86)
Market-to-book	-0.000***	-0.000	-0.000
	(-5.90)	(-0.56)	(-1.30)
Leverage	-1.084***	-0.518***	-0.602***
	(-8.32)	(-5.15)	(-6.17)
Complement − substitute			0.302
			(1.29)
Wald p			0.198
Cohort × Firm FE	Yes	Yes	Yes
Cohort × Year FE	Yes	Yes	Yes
Mean of dep. var.	20.194	151.208	20.194
Observations	26,718	41,238	42,446
Pseudo R^2	0.050	0.017	0.018

Table 13. The Effect of an Open-Source Release on Peer Firms' Financial Distress

This table reports stacked difference-in-differences estimates of the effect of open-source releases on peer firms' default risk. The dependent variable is the Merton distance to default in column (1) or the Ohlson O-score default probability in column (2). The Merton distance to default solves for the value and volatility of assets from the market value of equity, equity volatility over the calendar year and debt (current liabilities plus half of long-term debt) with a one-year horizon; higher values mean lower default risk, and firms without debt have no value. The Ohlson O-score default probability is the logistic transformation of the Ohlson (1980) O-score; higher values mean higher default risk. Both are winsorized at 1% and 99% within fiscal year. Table IA.11 reports other distress and restructuring measures. Post equals one from the fiscal year of the release onward. Complement and Substitute indicate the peer's group. Each column includes complement, substitute and control peers, with control peers carrying entropy-balancing weights that match them to complement and substitute peers together on the moments listed in Table 4, as in column (3) of Table 7. Complement − substitute is the difference between the two coefficients, with the p-value of a Wald test. Mean of dep. var. is the pre-release mean for complement peers. Control variables are size, market-to-book and leverage. All regressions include cohort-by-firm and cohort-by-year fixed effects, where each release defines a cohort. Standard errors are clustered at the release-by-fiscal-year level. The sample covers the 249 releases from 2010 to 2024 and, for each release, fiscal years $t-3$ to $t+3$. t-statistics are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Dependent variable =	Merton distance to default	Ohlson O-score PD
	(1)	(2)
Post × Complement	−0.166	−0.0157***
	(−1.38)	(−3.17)
Post × Substitute	−0.247***	0.0012
	(−3.96)	(0.57)
Complement − substitute	0.082	−0.0169***
	(0.68)	(−3.33)
Wald p	0.494	0.001
Controls	Yes	Yes
Cohort × Firm FE	Yes	Yes
Cohort × Year FE	Yes	Yes
Mean of dep. var.	7.974	0.2936
Observations	91,813	113,456
Adjusted R^2	0.811	0.823

Table 14. The
Effect of an Open-Source Release on Peer Firms' Financial Policy and Operating Income

This table reports stacked difference-in-differences estimates of the effect of open-source releases on peer firms' leverage, cash holdings, operating income and SG&A. The dependent variable is leverage (long-term plus current debt over total assets), cash (cash and short-term investments over total assets), operating income (operating income before depreciation over lagged total assets) or SG&A (selling, general and administrative expense over sales). Column (1) excludes leverage from the controls. Post equals one from the fiscal year of the release onward. Complement and Substitute indicate the peer's group. Each column includes complement, substitute and control peers, with control peers carrying entropy-balancing weights that match them to complement and substitute peers together on the moments listed in Table 4, as in column (3) of Table 7. Complement − substitute is the difference between the two coefficients, with the p-value of a Wald test. Mean of dep. var. is the pre-release mean for complement peers. Control variables are size, market-to-book and leverage. All regressions include cohort-by-firm and cohort-by-year fixed effects, where each release defines a cohort. Standard errors are clustered at the release-by-fiscal-year level. The sample covers the 249 releases from 2010 to 2024 and, for each release, fiscal years $t-3$ to $t+3$. t-statistics are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	Financial policy		Cost saving	
Dependent variable =	Leverage	Cash	Operating income	SG&A
	(1)	(2)	(3)	(4)
Post × Complement	−0.0148***	0.0094***	−0.0115***	0.0175***
	(−3.64)	(2.65)	(−3.72)	(2.73)
Post × Substitute	0.0271***	−0.0309***	0.0054***	−0.0034
	(16.02)	(−19.66)	(3.45)	(−1.55)
Complement − substitute	−0.0419***	0.0403***	−0.0168***	0.0209***
	(−10.18)	(10.67)	(−5.22)	(3.13)
Wald p	0.000	0.000	0.000	0.002
Controls	Yes	Yes	Yes	Yes
Cohort × Firm FE	Yes	Yes	Yes	Yes
Cohort × Year FE	Yes	Yes	Yes	Yes
Mean of dep. var.	0.2348	0.2829	0.0641	0.5118
Observations	116,971	116,971	116,924	115,431
Adjusted R^2	0.830	0.797	0.749	0.909

Internet Appendix for:
"Commoditize Your Complement: Who Bears the Cost of an Open-Source Release?"

A Appendix Tables

Table IA.1. Every Release in the Estimation Sample

This table lists the 249 releases in the estimation sample, grouped by releasing firm and sorted by date. Public is the date on which the repository became public. Peers is the number of firms in the release's peer pool before classification. Stars is the number of GitHub stars the repository has, and the count after each firm is the number of its releases. The list runs down the left half of the table and continues in the right half.

Repository	Public	Peers	Stars	Repository	Public	Peers	Stars
Advanced Micro Devices Inc (3)				**Microsoft Corp** (70)			
Xilinx/XRT	2018-10-03	280	685	mono/mono	2010-07-23	120	11.5k
Xilinx/Vitis-AI	2019-12-04	333	1,801	dotnet/reactive	2013-09-04	116	7,189
GPUOpen-LibrariesAndSDKs/ RadeonProRenderSDK	2020-05-14	375	285	dotnet/aspnetcore	2014-03-12	238	38.4k
Alphabet Inc (40)				microsoft/TypeScript	2014-06-18	187	110k
angular/angular.js	2010-01-07	80	58.5k	github/scientist	2014-11-18	180	7,754
google/lmctfy	2013-10-03	91	3,404	dotnet/fsharp	2015-01-12	280	4,325
kubernetes/kubernetes	2014-06-09	215	126k	dotnet/orleans	2015-01-26	213	10.9k
bazelbuild/bazel	2014-06-13	181	25.8k	dotnet/msbuild	2015-03-13	219	5,547
angular/material	2014-07-03	202	16.4k	microsoft/perfview	2015-03-30	210	4,741
google/leveldb	2014-08-28	288	39.4k	microsoft/WinObjC	2015-08-07	275	6,228
angular/angular	2014-09-30	191	101k	linkedin/PalDB	2015-09-30	310	939
google/personfinder	2014-10-20	181	548	microsoft/ Git-Credential-Manager-for-Windows	2015-10-01	234	2,939
golang/go	2014-12-09	209	138k	microsoft/code-push	2015-10-22	287	4,409
grpc/grpc	2015-02-27	244	45.3k	microsoft/CNTK	2015-11-27	312	17.6k
google/szl	2015-03-17	302	69	dotnet/vscode-csharp	2015-12-04	226	3,068
google/santa-tracker-web	2015-03-27	279	659	microsoft/react-native-windows	2015-12-16	207	17.3k
google/sagetv	2015-08-10	288	287	microsoft/pxt	2016-01-25	224	2,304
angular/components	2016-01-05	256	25.0k	microsoft/WSL	2016-04-07	257	33.6k
tensorflow/serving	2016-02-17	290	6,360	xamarin/Xamarin.Forms	2016-04-28	226	5,534
openthread/openthread	2016-05-12	305	4,025	microsoft/XmlNotepad	2016-04-28	259	1,192
google/nomulus	2016-10-19	247	1,807	microsoft/fluentui	2016-06-07	235	20.2k
google-deepmind/lab	2016-12-01	302	7,373	PowerShell/PowerShell	2016-08-19	227	55.2k
google/grumpy	2017-01-05	321	10.5k	microsoft/GraphEngine	2016-09-06	285	2,263
google/earthenterprise	2017-03-27	356	2,749	microsoft/ExaStore	2016-11-17	332	35
google-deepmind/sonnet	2017-04-04	304	9,966	microsoft/VFSForGit	2017-01-26	274	6,136
tensorflow/tfjs	2018-03-06	357	19.1k	microsoft/AirSim	2017-02-15	324	18.4k
google/ground-android	2018-04-03	273	293	microsoft/napajs	2017-08-09	339	9,173
google/ground-platform	2018-04-26	283	249	microsoft/terminal	2017-08-14	288	104k
google/go-cloud	2018-07-24	268	9,911	microsoft/workbooks	2017-09-07	251	463
google/starlark-go	2018-10-24	262	2,758	microsoft/pai	2017-09-26	293	2,688
google-research/bert	2018-10-26	288	40.0k	Azure/AKS	2017-10-26	284	2,138
jax-ml/jax	2018-10-26	354	36.3k	microsoft/azuredatastudio	2017-11-03	253	7,684
tensorflow/tfx	2019-02-05	288	2,192	microsoft/fast	2018-03-05	284	9,669
google-ai-edge/mediapipe	2019-06-14	341	36.9k	microsoft/service-fabric	2018-03-15	364	3,065
GoogleCloudPlatform/iap-desktop	2019-06-27	313	878	microsoft/winfile	2018-04-03	278	7,063
google/OpenSK	2020-01-31	405	3,420	microsoft/Detours	2018-04-16	293	6,369
tensorflow/quantum	2020-03-10	406	2,180	dotnet/machinelearning	2018-05-08	266	9,351
google-deepmind/alphafold	2021-06-18	370	14.8k	microsoft/MS-DOS	2018-06-26	304	32.3k
google-deepmind/mujoco	2021-08-30	354	15.0k	Azure/iotedge	2018-06-28	370	1,508
google/rune	2021-12-06	372	1,929	microsoft/microsoft-ui-xaml	2018-07-27	282	8,306
google/magika	2024-02-16	477	18.5k	github/codeql	2018-08-01	281	10.1k
google-deepmind/gemma	2024-02-21	467	5,717	microsoft/nni	2018-09-10	353	14.4k
google-ai-edge/LiteRT	2024-09-05	492	3,380	microsoft/Trill	2018-09-27	331	1,271
google-deepmind/alphafold3	2024-11-12	493	8,530	dotnet/winforms	2018-10-22	279	4,852
Amazon.com Inc (15)				dotnet/wpf	2018-10-22	277	7,722
aws/s2n-tls	2014-06-30	238	4,761	microsoft/onnxruntime	2018-12-04	344	21.7k

continued

Table IA.1, continued

Repository	Public	Peers	Stars
aws/lumberyard	2017-08-16	286	2,137
corretto/corretto-8	2018-11-15	327	2,124
corretto/corretto-11	2019-02-12	345	674
bottlerocket-os/bottlerocket	2020-03-11	352	9,668
aws-cloudformation/cloudformation-guard	2020-06-17	357	1,386
opensearch-project/OpenSearch	2021-04-13	415	13.6k
opensearch-project/OpenSearch-Dashboards	2021-04-13	435	2,119
aws-amplify/amplify-ui	2021-06-03	355	1,142
corretto/corretto-17	2021-08-10	361	247
aws/eks-anywhere	2021-09-09	394	2,098
corretto/corretto-18	2022-02-01	408	19
corretto/corretto-19	2022-08-12	407	11
corretto/corretto-20	2022-12-16	429	16
awslabs/llrt	2023-09-21	507	8,799
Apple Inc (2)			
apple/foundationdb	2018-04-20	280	16.7k
apple/security-pcc	2024-10-25	397	1,046
Atlassian Corp Plc (1)			
atlassian/pragmatic-drag-and-drop	2024-02-13	400	12.8k
Autodesk Inc (1)			
Autodesk/hig	2017-04-04	333	193
Bentley Systems Inc (2)			
iTwin/iTwinUI-react	2021-04-16	362	84
iTwin/iTwinUI	2021-04-16	373	124
Blackbaud Inc (2)			
blackbaud/skyux2	2016-05-12	329	30
blackbaud/skyux	2022-02-28	413	55
Broadcom Inc (1)			
vmware/splinterdb	2022-03-21	486	732
Cisco Systems Inc (2)			
threatgrid/asami	2018-04-10	295	662
momentum-design/momentum-ui	2019-01-07	335	212
Cloudflare Inc (1)			
cloudflare/workerd	2022-09-28	451	8,671
Crowdstrike Holdings Inc (1)			
CrowdStrike/ember-toucan-core	2023-02-02	413	11
Datadog Inc (1)			
vectordotdev/vector	2018-08-28	299	22.5k
Dynatrace Inc (1)			
dynatrace-oss/barista	2019-10-31	320	146
Electronic Arts Inc (2)			
electronicarts/CnC_Remastered_Collection	2020-06-02	302	21.4k
electronicarts/CNC_TS_and_RA2_Mission_Editor	2024-03-13	419	184
Epam Systems Inc (3)			
epam/ketcher	2017-09-25	314	879
epam/fix-antenna-net-core	2021-12-27	380	49
epam/ai-dial-core	2023-09-25	448	658
Intel Corp (6)			
intelsdi-x/snap	2015-12-03	321	1,780
GameTechDev/PresentMon	2016-03-10	339	2,542
intel/zephyr.js	2016-06-21	324	182
intel/pcm	2016-11-01	314	3,322
intel/ref-os-iot	2017-02-22	356	25
openvinotoolkit/openvino	2018-10-16	369	10.8k
International Business Machines (6)			
ansible/ansible	2012-03-07	87	70.6k
carbon-design-system/carbon	2017-03-14	343	9,412
ansible/awx	2017-09-08	297	15.5k
carbon-design-system/carbon-components-angular	2018-08-01	278	564
microsoft/accessibility-insights-windows	2018-12-10	262	546
microsoft/calculator	2019-01-29	337	31.0k
AzureAD/microsoft-authentication-library-for-java	2019-03-05	318	340
Azure/Industrial-IoT	2019-03-05	304	576
microsoft/PowerToys	2019-05-07	293	138k
microsoft/BotFramework-Composer	2019-11-05	319	891
dotnet/runtime	2019-11-27	334	18.2k
microsoft/WindowsAppSDK	2020-04-16	321	4,670
microsoft/clarity	2020-05-01	330	2,728
microsoft/azurelinux	2020-07-23	309	5,338
Azure/bicep	2020-08-31	309	3,636
microsoft/onefuzz	2020-09-16	411	2,824
microsoft/Extensible-Storage-Engine	2021-02-01	346	849
linkedin/venice	2021-03-19	375	611
microsoft/go	2022-02-09	410	430
Clean-Reader/CleanReader.Desktop	2022-04-11	229	400
microsoft/Microsoft-3D-Movie-Maker	2022-05-05	425	3,835
Azure/data-api-builder	2023-03-16	392	1,506
PowerShell/AIShell	2023-04-03	418	517
PowerShell/DSC	2023-08-08	417	515
microsoft/sudo	2024-02-09	389	5,859
microsoft/garnet	2024-03-19	512	12.0k
microsoft/openvmm	2024-10-17	383	1,884
microsoft/BitNet	2024-10-18	393	40.2k
Mongodb Inc (1)			
mongodb/leafygreen-ui	2018-11-09	268	271
Morgan Stanley Dean Witter & Co (1)			
morganstanley/hobbes	2017-06-16	363	1,227
National Instruments Corp (2)			
ni/design-system	2017-05-31	335	6
ni/nimble	2022-09-21	398	39
Netapp Inc (1)			
NetApp/trident	2016-12-22	327	876
Netflix Inc (7)			
Netflix/ribbon	2013-01-29	99	4,616
Netflix/genie	2013-06-24	99	1,766
Netflix/security_monkey	2014-07-01	229	4,371
Netflix/atlas	2014-12-15	249	3,564
Netflix/falcor	2015-08-18	318	10.6k
spinnaker/spinnaker	2015-11-17	245	9,782
Netflix/iceberg	2018-01-05	310	494
Nvidia Corp (9)			
NVIDIA/aistore	2017-12-15	305	1,921
NVIDIA/DALI	2018-06-20	366	5,740
triton-inference-server/server	2018-10-05	301	11.0k
NVIDIA/NVFlare	2021-08-06	366	963
NVIDIA/DLSS	2021-08-25	452	1,440
NVIDIAGameWorks/NVIDIAImageScaling	2021-11-17	420	615
NVlabs/instant-ngp	2022-01-06	497	17.5k
NVIDIA/cuda-quantum	2023-03-22	525	1,124
NVIDIA/TensorRT-LLM	2023-10-20	508	14.5k
Oracle Corp (27)			
oracle/opengrok	2013-01-08	86	4,929
mysql/mysql-server	2014-09-29	217	12.4k
mysql/mysql-workbench	2015-10-26	242	1,001
oracle/graal	2016-01-15	290	21.7k
oracle/pgql-lang	2016-06-22	241	205
oracle/railcar	2017-06-30	327	1,120
openjdk/jdk	2018-09-18	294	23.3k

continued

Table IA.1, continued

Repository	Public	Peers	Stars
carbon-design-system/carbon-web-components	2019-05-16	371	474
openshift/microshift	2021-04-27	468	841
Mastercard Inc (1)			
Mastercard/made	2023-01-09	482	19
Meta Platforms Inc (19)			
facebook/hhvm	2010-01-04	83	18.7k
facebook/proxygen	2014-11-06	219	8,379
facebook/openbmc	2015-03-11	315	687
facebook/fboss	2015-03-11	260	989
facebook/sapling	2016-05-06	259	7,008
Instagram/IGListKit	2016-10-12	296	13.1k
pytorch/pytorch	2017-01-19	359	102k
facebook/flipper	2018-04-13	314	13.5k
facebookincubator/katran	2018-05-23	309	5,307
facebookarchive/LogDevice	2018-09-13	278	1,900
facebook/Rapid	2018-12-11	297	639
facebookresearch/nevergrad	2018-12-21	340	4,205
facebookresearch/habitat-lab	2019-02-05	313	3,118
facebookresearch/habitat-sim	2019-02-22	385	3,807
facebook/Ax	2019-05-02	287	2,798
pytorch/serve	2020-04-22	393	4,344
facebookresearch/fairseq2	2022-12-23	431	1,145
meta-llama/llama	2023-02-15	400	59.6k
meta-llama/llama3	2024-04-19	399	29.2k
openjdk/jdk12u	2019-04-22	309	16
openjdk/jdk11u	2019-04-24	302	163
openjdk/jfx	2019-05-07	324	3,286
graalvm/labs-openjdk-11	2019-10-01	362	44
openjdk/jdk14	2019-12-19	309	16
oracle/coherence	2020-06-22	434	470
graalvm/labs-openjdk-15	2020-09-15	391	4
openjdk/nashorn	2020-11-06	334	495
openjdk/jdk16u	2020-12-21	336	24
graalvm/labs-openjdk-16	2021-02-08	347	2
openjdk/jdk8u	2021-02-17	360	305
openjdk/jdk17u	2021-06-17	363	127
openjdk/jfx17u	2021-09-13	335	6
openjdk/jdk18u	2021-12-15	369	6
openjdk/jdk19u	2022-06-17	418	19
openjdk/jdk21u	2023-06-16	418	99
openjdk/jfx21u	2023-07-20	389	10
graalvm/graalvm-community-jdk21u	2023-11-28	470	17
openjdk/jfx23u	2024-07-12	340	2
oracle/sandwood	2024-08-06	471	36
Paypal Holdings Inc (1)			
paypal/junodb	2023-04-12	455	2,635
Pier 1 Imports Inc (1)			
Pier1/rocketbelt	2016-10-13	308	11
Ping Identity Holding Corp (1)			
pingidentity/astro	2021-12-30	362	35
Red Hat Inc (1)			
3scale/porta	2018-09-20	364	78
Salesforce Inc (8)			
forcedotcom/dataloader	2011-11-07	84	541
salesforce-ux/design-system	2015-09-02	304	3,701
salesforce/design-system-react	2017-10-19	262	981
mulesoft/mule	2019-08-01	352	500
slackapi/bolt-python	2020-08-10	319	1,323
SalesforceCommerceCloud/pwa-kit	2021-09-09	429	359
SalesforceLabs/ampscript-core	2023-06-13	488	32
SalesforceFoundation/NPSP	2024-10-17	431	44
Schrodinger Inc (1)			
schrodinger/sketcher	2023-07-27	476	40
Target Corp (1)			
go-vela/server	2019-11-05	307	118
Teradata Corp De (1)			
Teradata/kylo	2017-02-09	298	1,113
Twilio Inc (1)			
twilio-labs/paste	2019-08-09	302	478
Tyler Corp (1)			
tyler-technologies-oss/forge	2022-08-18	466	68
Uber Technologies Inc (2)			
uber/baseweb	2018-03-12	274	9,004
uber-go/mock	2023-05-04	515	3,404
Workday Inc (1)			
Workday/canvas-kit	2019-07-03	309	358
Workiva Inc (1)			
Workiva/eva	2019-05-30	288	561

Table IA.2. The Releasing Firms

This table lists the 41 releasing firms, sorted by number of releases. Years is the range of release years. Assets, sales and R&D are medians over the firm's fiscal years of release, in billions of dollars. Peers is the median size of the peer pool across the firm's releases.

Releasing firm	Releases	Years	Assets	Sales	R&D	Peers
Microsoft Corp	70	2010–2024	258.8	110.4	14.73	290
Alphabet Inc	40	2010–2024	197.3	110.9	16.62	289
Oracle Corp	27	2013–2024	123.3	39.7	6.12	336
Meta Platforms Inc	19	2010–2024	115.4	63.3	11.94	313
Amazon.com Inc	15	2014–2023	273.2	333.3	39.34	361
Nvidia Corp	9	2017–2023	41.2	26.9	5.27	420
Salesforce Inc	8	2011–2024	60.7	19.2	3.18	335
Netflix Inc	7	2013–2018	8.6	6.1	0.56	245
Intel Corp	6	2015–2018	118.3	61.1	12.92	331
International Business Machines	6	2012–2021	125.4	79.1	5.99	320
Advanced Micro Devices Inc	3	2018–2020	6.0	6.7	1.55	333
Epam Systems Inc	3	2017–2023	3.5	3.8	0.00	380
Apple Inc	2	2018–2024	365.4	328.2	22.80	338
Bentley Systems Inc	2	2021	2.7	1.0	0.22	367
Blackbaud Inc	2	2016–2022	2.2	0.9	0.12	371
Cisco Systems Inc	2	2018–2019	103.3	50.6	6.45	315
Electronic Arts Inc	2	2020–2024	12.8	6.5	2.17	360
National Instruments Corp	2	2017–2022	2.0	1.5	0.28	366
Uber Technologies Inc	2	2018–2023	38.7	37.3	3.16	394
Atlassian Corp Plc	1	2024	5.2	4.4	2.18	400
Autodesk Inc	1	2017	4.1	2.1	0.76	333
Broadcom Inc	1	2022	73.2	33.2	4.92	486
Cloudflare Inc	1	2022	2.6	1.0	0.30	451
Crowdstrike Holdings Inc	1	2023	6.6	3.1	0.77	413
Datadog Inc	1	2018	–	–	–	299
Dynatrace Inc	1	2019	2.0	0.5	0.12	320
Mastercard Inc	1	2023	42.4	25.1	0.00	482
Mongodb Inc	1	2018	0.7	0.3	0.09	268
Morgan Stanley Dean Witter & Co	1	2017	851.7	43.6	0.00	363
Netapp Inc	1	2016	9.5	5.5	0.78	327
Paypal Holdings Inc	1	2023	82.2	29.8	1.60	455
Pier 1 Imports Inc	1	2016	0.8	1.8	0.00	308
Ping Identity Holding Corp	1	2021	1.2	0.3	0.08	362
Red Hat Inc	1	2018	5.6	3.4	0.67	364
Schrodinger Inc	1	2023	0.8	0.2	0.18	476
Target Corp	1	2019	42.8	78.1	0.00	307
Teradata Corp De	1	2017	2.6	2.2	0.31	298
Twilio Inc	1	2019	5.2	1.1	0.39	302
Tyler Corp	1	2022	4.7	1.9	0.11	466
Workday Inc	1	2019	6.8	3.6	1.55	309
Workiva Inc	1	2019	0.6	0.3	0.09	288
Total	249	2010–2024				

Table IA.3. Complement and Substitute Peers, Release by Release

This table lists the complement and substitute peers the classifier assigns to each of the 249 releases in the estimation sample, so that the classification can be checked against knowledge of these firms rather than against a summary statistic alone. Releases are sorted by releasing firm and then by date. *N* is the number of peers in the group, and the names beside it are the three ranked highest on the technology axis. Three names cover the whole group for 57 percent of the complement groups, whose median size is three, against a median of 21 for the substitute groups. Company names are Compustat `conm` with the legal suffix removed; the date is the release year and month; the repository is shown without its owner prefix, which is the releasing firm in every row. A dash indicates that the release has no peer in that group. The control group is counted in Table 4 but not named here, as it runs to 526 firms.

			Complement peers		Substitute peers	
Firm	Date	Repository	*N*	Highest on technology	*N*	Highest on technology
Advanced Micro Devices	2018-10	`XRT`	3	Asure Software, Diebold Nixdorf, Blackbaud	22	Mellanox Technologies, NVIDIA, Ambarella
	2019-12	`Vitis-AI`	5	Blackbaud, Paycom Software, Ringcentral	38	NVIDIA, Silicon Laboratories, Mellanox Technologies
	2020-05	`RadeonProRenderSDK`	8	Dynatrace, Intralinks Holdings, Nexttrip	24	Autodesk, Meta Platforms, Citrix Systems
Alphabet	2010-01	`angular.js`	2	Semtech, Syntel	8	PTC, Intl Business Machines, Autodesk
	2013-10	`lmctfy`	5	Syntel, Semtech, Computer Task Group	11	Intl Business Machines, Citrix Systems, CA
	2014-06	`kubernetes`	4	Nimble Storage, Virtusa, Micros Systems	13	Intl Business Machines, Open Text, Oracle
	2014-06	`bazel`	1	Sajan	11	Intl Business Machines, Open Text, Oracle
	2014-07	`material`	3	Infoblox, Ipass, Dss	18	Liveone, Open Text, Autodesk
	2014-08	`leveldb`	5	Liveramp Holdings, Carbonite, Factset Research Systems	29	Seagate Technology Holdings, Western Digital, Red Hat
	2014-09	`angular`	3	Liveramp Holdings, Digi International, Servicesource Intl	7	Citrix Systems, Autodesk, Intuit
	2014-10	`personfinder`	4	Servicenow, Pegasystems, Epam Systems	3	Meta Platforms, Autoweb, Open Text
	2014-12	`go`	4	Perficient, Virtusa, KeyW Holding	9	Brightcove, Oracle, Intuit
	2015-02	`grpc`	2	Infoblox, Netsuite	13	Salesforce, Red Hat, PTC
	2015-03	`szl`	2	Spansion, Computer Task Group	12	Intl Business Machines, CA, Adobe
	2015-03	`santa-tracker-web`	4	Mellanox Technologies, Widepoint, Spansion	26	Verisign, Box, Open Text
	2015-08	`sagetv`	4	Mellanox Technologies, Infoblox, Spansion	25	Smith Micro Software, PTC, Red Hat
	2016-01	`components`	2	Infoblox, Activision Blizzard	9	Autodesk, Open Text, Earthlink Holdings
	2016-02	`serving`	4	Liveramp Holdings, Pixelworks, Netlist	18	Synopsys, Adobe, PTC
	2016-05	`openthread`	10	Computer Task Group, Mobileiron, Angie's List	17	Synopsys, Citrix Systems, PTC
	2016-10	`nomulus`	6	Infoblox, Servicesource Intl, Code Rebel	18	Verisign, Web.com Group, Earthlink Holdings
	2016-12	`lab`	4	Netlist, Spansion, Cypress Semiconductor	21	Intl Business Machines, Microsoft, Adobe
	2017-01	`grumpy`	8	Cypress Semiconductor, Mellanox Technologies, Spansion	25	Intl Business Machines, PTC, Salesforce
	2017-03	`earthenterprise`	5	Kratos Defense & Security, Aerohive Networks, F5	10	Open Text, Box, Intl Business Machines
	2017-04	`sonnet`	4	Liveramp Holdings, Netlist, Cypress Semiconductor	20	Intl Business Machines, Open Text, Autodesk
	2018-03	`tfjs`	5	Spansion, Mellanox Technologies, Cypress Semiconductor	22	Altair Engineering, Citrix Systems, Autodesk
	2018-04	`ground-android`	1	PDF Solutions	7	Open Text, Linkedin, Ziff Davis
	2018-04	`ground-platform`	6	Infoblox, Fortinet, Medidata Solutions	18	Box, Yext, Earthlink Holdings
	2018-07	`go-cloud`	5	Virtusa, Intralinks Holdings, Computer Task Group	15	Linkedin, Open Text, Oracle
	2018-10	`starlark-go`	2	A10 Networks, Liveramp Holdings	19	Red Hat, PTC, Salesforce

			Complement peers		Substitute peers	
Firm	Date	Repository	*N*	Highest on technology	*N*	Highest on technology
	2018-10	`bert`	2	Simulations Plus, Falconstor Software	11	Oracle, Open Text, Autodesk
	2018-10	`jax`	3	Computer Task Group, Okta, Virtusa	24	Intl Business Machines, Citrix Systems, Red Hat
	2019-02	`tfx`	3	Computer Task Group, Virtusa, A10 Networks	20	PTC, Intl Business Machines, Open Text
	2019-06	`mediapipe`	3	Mellanox Technologies, Pixelworks, Cypress Semiconductor	22	Meta Platforms, Dropbox, Logmein
	2019-06	`iap-desktop`	9	Infoblox, Code Rebel, Ipass	22	Citrix Systems, Box, Logmein
	2020-01	`OpenSK`	11	Ping Identity Holding, Netscout Systems, Arista Networks	29	Synopsys, Semtech, NVIDIA
	2020-03	`quantum`	3	Virtusa, Computer Task Group, Intralinks Holdings	33	NVIDIA, Silicon Laboratories, Synopsys
	2021-06	`alphafold`	6	Simulations Plus, Fortinet, Medidata Solutions	33	Intl Business Machines, Box, Salesforce
	2021-08	`mujoco`	9	Alteryx, Matterport, Okta	31	Altair Engineering, Adobe, Microsoft
	2021-12	`rune`	3	Cornerstone Ondemand, Fastly, Infoblox	28	Intl Business Machines, Red Hat, Salesforce
	2024-02	`magika`	4	Infoblox, Liveramp Holdings, Netscout Systems	25	Open Text, Brightcove, Citrix Systems
	2024-02	`gemma`	1	Liveramp Holdings	13	Intl Business Machines, Intel, Citrix Systems
	2024-09	`LiteRT`	16	Fastly, Audioeye, New Relic	39	NVIDIA, Citrix Systems, Altair Engineering
	2024-11	`alphafold3`	10	Rapid7, Fleetmatics Group, Pegasystems	22	Godaddy, Meta Platforms, Open Text
Amazon.com	2014-06	`s2n-tls`	9	Radisys, Mandiant, Kratos Defense & Security	19	CA, Intl Business Machines, Citrix Systems
	2017-08	`lumberyard`	5	Intralinks Holdings, GTY Technology Holdings, Diebold Nixdorf	24	Microsoft, Citrix Systems, Electronic Arts
	2018-11	`corretto-8`	3	Diebold Nixdorf, Radisys, GTY Technology Holdings	22	Citrix Systems, Verisign, Microsoft
	2019-02	`corretto-11`	4	Diebold Nixdorf, Radisys, Smartsheet	22	Citrix Systems, Verisign, Microsoft
	2020-03	`bottlerocket`	11	Kratos Defense & Security, Pluralsight, Ping Identity Holding	25	Intl Business Machines, Box, Citrix Systems
	2020-06	`cloudformation-guard`	12	Kratos Defense & Security, Cambium Learning Group, Virtusa	37	Box, Red Hat, Godaddy
	2021-04	`OpenSearch`	3	Computer Task Group, Widepoint, GTY Technology Holdings	30	Rightside Group, Pinterest, Linkedin
	2021-04	`OpenSearch-Dashboards`	3	Computer Task Group, Exela Technologies, PDF Solutions	28	Yext, Cargurus, Linkedin
	2021-06	`amplify-ui`	12	Zscaler, Computer Task Group, Virtusa	31	Citrix Systems, Intl Business Machines, Earthlink Holdings
	2021-08	`corretto-17`	3	Diebold Nixdorf, New Relic, Audioeye	12	Broadsoft, Microsoft, Open Text
	2021-09	`eks-anywhere`	4	Kratos Defense & Security, Virtusa, Computer Task Group	16	Intl Business Machines, Open Text, Oracle
	2022-02	`corretto-18`	2	Diebold Nixdorf, Cornerstone Ondemand	9	Uipath, Intl Business Machines, Microsoft
	2022-08	`corretto-19`	3	Diebold Nixdorf, Audioeye, Radisys	13	Uipath, Veritone, Intl Business Machines
	2022-12	`corretto-20`	3	Gitlab, Diebold Nixdorf, Intellicheck	17	Open Text, Intl Business Machines, Microsoft
	2023-09	`llrt`	6	Kratos Defense & Security, Coherent, Light & Wonder	42	Citrix Systems, Twilio, Box
Apple	2018-04	`foundationdb`	8	Alteryx, Sajan, Super Micro Computer	23	Red Hat, Open Text, Salesforce
	2024-10	`security-pcc`	2	Enernoc, Diebold Nixdorf	15	Twilio, Informatica, Docusign
Atlassian	2024-02	`pragmatic-drag-and-drop`	4	Dynatrace, Blend Labs, Bigbear.ai Holdings	25	Meta Platforms, Autodesk, New Relic
Autodesk	2017-04	`hig`	2	Light & Wonder, NCI	31	Box, Intuit, Adobe

			Complement peers		Substitute peers	
Firm	Date	Repository	*N*	Highest on technology	*N*	Highest on technology
Bentley Systems	2021-04	iTwinUI	6	Trubridge, Cavium, Semtech	37	Cadence Design Systems, Intl Business Machines, Red Hat
	2021-04	iTwinUI-react	4	Cavium, Trubridge, Light & Wonder	7	Autodesk, Meta Platforms, Pluralsight
Blackbaud	2016-05	skyux2	25	Invensense, Pixelworks, Semtech	37	Box, Synacor, Adobe
	2022-02	skyux	1	Light & Wonder	17	Logmein, Super League Enterprise, Nextdoor Holdings
Broadcom	2022-03	splinterdb	4	GTY Technology Holdings, Cvent Holding, Datto Holding	29	CA, Red Hat, Datawatch
Cisco Systems	2018-04	asami	5	Corelogic, Liveramp Holdings, Stamps.com	21	Intl Business Machines, Intuit, 8x8
	2019-01	momentum-ui	1	Exela Technologies	23	Adobe, Autodesk, Intl Business Machines
Cloudflare	2022-09	workerd	1	Vertex	34	Citrix Systems, Intl Business Machines, Red Hat
Crowdstrike Holdings	2023-02	ember-toucan-core	4	TiVo, Semtech, Penguin Solutions	25	Dropbox, Atlassian, Intl Business Machines
Datadog	2018-08	vector	8	Semtech, Stamps.com, PDF Solutions	21	Commvault Systems, Citrix Systems, Intl Business Machines
Dynatrace	2019-10	barista	8	Textura, Semtech, Rimini Street	15	Intl Business Machines, Solarwinds, Open Text
Electronic Arts	2020-06	CnC_Remastered_Collection	4	Dynatrace, Blacksky Technology, Rekor Systems	10	Microsoft, Citrix Systems, Salesforce
	2024-03	CNC_TS_and_RA2_Mission_Editor	10	Nutanix, Blackline, Intralinks Holdings	38	Citrix Systems, Autodesk, Red Hat
Epam Systems	2017-09	ketcher	18	Linear Technology, Power Integrations, Pixelworks	23	Intuit, Verisign, Autodesk
	2021-12	fix-antenna-net-core	3	Ping Identity Holding, Semtech, Diebold Nixdorf	26	Citrix Systems, Salesforce, Red Hat
	2023-09	ai-dial-core	6	Liveramp Holdings, Electronics for Imaging, Exela Technologies	40	Earthlink Holdings, Atlassian, Meta Platforms
Intel	2015-12	snap	3	Qualstar, Corelogic, Coherent	25	Red Hat, Intl Business Machines, Open Text
	2016-03	PresentMon	3	Coherent, NCI, Servicesource Intl	20	Salesforce, Citrix Systems, Intl Business Machines
	2016-06	zephyr.js	3	Light & Wonder, Intralinks Holdings, Corelogic	30	Citrix Systems, Advanced Micro Devices, Microsoft
	2016-11	pcm	5	Enernoc, Diebold Nixdorf, Corelogic	33	Advanced Micro Devices, Hewlett Packard Enterprise, Spansion
	2017-02	ref-os-iot	1	Pegasystems	22	Intl Business Machines, Digi International, Sigma Designs
	2018-10	openvino	2	Rekor Systems, Diebold Nixdorf	26	Intl Business Machines, Citrix Systems, Red Hat
Intl Business Machines	2012-03	ansible	2	Diebold Nixdorf, Semtech	12	PTC, Citrix Systems, CA
	2017-03	carbon	2	Ixys, Light & Wonder	16	Synopsys, Adobe, Box
	2017-09	awx	2	TiVo, Enernoc	13	Box, Adobe, Meta Platforms
	2018-08	carbon-components-angular	1	Liveramp Holdings	17	Cloudera, Box, Meta Platforms
	2019-05	carbon-web-components	8	Invensense, Light & Wonder, Coherent	34	PTC, Red Hat, Box
	2021-04	microshift	3	Pixelworks, Wolfspeed, Coherent	28	Nutanix, Rackspace Technology, Commvault Systems
Mastercard	2023-01	made	4	Light & Wonder, Black Box, Liveramp Holdings	31	Box, Intl Business Machines, Citrix Systems
Meta Platforms	2010-01	hhvm	1	Micros Systems	11	Intl Business Machines, Citrix Systems, Adobe
	2014-11	proxygen	1	Dss	21	Verisign, Intuit, Autodesk
	2015-03	fboss	4	Gigamon, Inovalon Holdings, CSP	31	Juniper Networks, Red Hat, Aruba Networks
	2015-03	openbmc	1	Roper Technologies	12	Silicon Graphics Intl, Intl Business Machines, Intel

			Complement peers		Substitute peers	
Firm	Date	Repository	*N*	Highest on technology	*N*	Highest on technology
	2016-05	sapling	3	Cleanspark, Light & Wonder, Diebold Nixdorf	23	Citrix Systems, Salesforce, Intl Business Machines
	2016-10	IGListKit	2	Qualys, Astea International	16	8x8, Upland Software, Web.com Group
	2017-01	pytorch	1	Rimini Street	25	Intl Business Machines, Intel, Red Hat
	2018-04	flipper	3	GTY Technology Holdings, Rimini Street, Optimizerx	21	Intl Business Machines, Earthlink Holdings, Cloudera
	2018-05	katran	5	Zscaler, Code Rebel, Appfolio	16	Citrix Systems, Juniper Networks, Ribbon Communications
	2018-09	LogDevice	3	Enernoc, I3 Verticals, PDF Solutions	17	Intl Business Machines, Citrix Systems, Box
	2018-12	Rapid	3	Aware, CSP, Trubridge	13	Box, Open Text, Intuit
	2018-12	nevergrad	1	Trubridge	17	Intl Business Machines, Verint Systems, Altair Engineering
	2019-02	habitat-lab	2	Par Technology, CSP	28	Red Hat, Intl Business Machines, PTC
	2019-02	habitat-sim	1	Par Technology	35	Altair Engineering, Intl Business Machines, Red Hat
	2019-05	Ax	5	Nutanix, Coupa Software, Rekor Systems	19	Citrix Systems, Red Hat, Salesforce
	2020-04	serve	1	Simulations Plus	33	Reachlocal, Open Text, Earthlink Holdings
	2022-12	fairseq2	3	Talend, Roper Technologies, Zuora	26	NVIDIA, Intl Business Machines, Citrix Systems
	2023-02	llama	7	Arteris, Simulations Plus, Creative Realities	39	Atlassian, Verisign, Box
	2024-04	llama3	4	Dynatrace, Simulations Plus, GTY Technology Holdings	36	Atlassian, Verisign, Open Text
Microsoft	2010-07	mono	2	Qualstar, Factset Research Systems	13	Intl Business Machines, PTC, Synopsys
	2013-09	reactive	2	Coherent, Light & Wonder	15	PTC, Gen Digital, Intl Business Machines
	2014-03	aspnetcore	1	Widepoint	21	Salesforce, Citrix Systems, Broadsoft
	2014-06	TypeScript	2	Semtech, Medidata Solutions	5	Salesforce, Intuit, Citrix Systems
	2014-11	scientist	1	Semtech	10	Earthlink Holdings, Web.com Group, Ziff Davis
	2015-01	fsharp	4	Kimball Electronics, Amkor Technology, Cypress Semiconductor	20	Salesforce, Synopsys, Adobe
	2015-01	orleans	4	Semtech, Liveramp Holdings, Light & Wonder	24	Box, 8x8, Commvault Systems
	2015-03	msbuild	2	Semtech, Intralinks Holdings	16	Box, Autodesk, Open Text
	2015-03	perfview	3	Semtech, Roper Technologies, Intralinks Holdings	20	Salesforce, Red Hat, Citrix Systems
	2015-08	WinObjC	2	Pixelworks, Semtech	18	Salesforce, Citrix Systems, Intuit
	2015-09	PalDB	3	Factset Research Systems, New Relic, Liveramp Holdings	27	Smith Micro Software, Salesforce, Intl Business Machines
	2015-10	Git-Credential-Manager-for-Windows	5	Intralinks Holdings, Cleanspark, Semtech	18	Red Hat, Salesforce, Citrix Systems
	2015-10	code-push	4	Mellanox Technologies, Widepoint, Spansion	30	Red Hat, Citrix Systems, PTC
	2015-11	CNTK	3	Liveramp Holdings, NCI, Widepoint	27	Intl Business Machines, Citrix Systems, Box
	2015-12	vscode-csharp	2	Semtech, Roper Technologies	8	Oracle, Autodesk, Open Text
	2015-12	react-native-windows	1	Semtech	10	Autodesk, Earthlink Holdings, Synacor
	2016-01	pxt	4	Intralinks Holdings, Semtech, Stamps.com	15	Box, Citrix Systems, Open Text
	2016-04	WSL	4	Super Micro Computer, Semtech, Medidata Solutions	10	Open Text, Egain, Earthlink Holdings
	2016-04	XmlNotepad	1	Intralinks Holdings	13	Open Text, Citrix Systems, Intuit
	2016-04	Xamarin.Forms	3	Intralinks Holdings, Semtech, Servicesource Intl	14	Synacor, Earthlink Holdings, Web.com Group
	2016-06	fluentui	5	Stamps.com, Bazaarvoice, Sajan	12	Commvault Systems, Intl Business Machines, Adobe
	2016-08	PowerShell	2	Liveramp Holdings, Medidata Solutions	15	Salesforce, Citrix Systems, Intuit
	2016-09	GraphEngine	5	Spansion, Netlist, Cypress Semiconductor	13	Open Text, Intl Business Machines, Verisign

Firm	Date	Repository	Complement peers		Substitute peers	
			N	Highest on technology	*N*	Highest on technology
	2016-11	ExaStore	1	Coherent	26	Netapp, Open Text, NVIDIA
	2017-01	VFSForGit	2	Semtech, Everpure	13	Open Text, Earthlink Holdings, Verisign
	2017-02	AirSim	10	Ambarella, Coherent, GSI Technology	27	Guidance Software, Synopsys, Citrix Systems
	2017-08	napajs	1	Liveramp Holdings	16	Salesforce, Intl Business Machines, Red Hat
	2017-08	terminal	4	Semtech, Outerwall, GTY Technology Holdings	28	PTC, Citrix Systems, Open Text
	2017-09	workbooks	3	Cambium Learning Group, Liveramp Holdings, Semtech	14	Verisign, Open Text, Meta Platforms
	2017-09	pai	4	Semtech, Qlogic, NCI	21	Commvault Systems, Box, Intl Business Machines
	2017-10	AKS	3	Liveramp Holdings, Servicesource Intl, Rimini Street	22	Open Text, Commvault Systems, Intuit
	2017-11	azuredatastudio	3	Semtech, Pixelworks, Liveramp Holdings	21	Intuit, Commvault Systems, Box
	2018-03	fast	2	Semtech, NCI	17	Commvault Systems, Meet Group Inc (the), Meta Platforms
	2018-03	service-fabric	8	Mellanox Technologies, Super Micro Computer, Spansion	36	Red Hat, Intl Business Machines, Citrix Systems
	2018-04	winfile	1	GTY Technology Holdings	13	Intuit, Citrix Systems, Red Hat
	2018-04	Detours	1	Semtech	18	Box, Open Text, Upland Software
	2018-05	machinelearning	3	Liveramp Holdings, Diebold Nixdorf, CSP	18	Box, Citrix Systems, Open Text
	2018-06	MS-DOS	1	Semtech	15	Box, Godaddy, Citrix Systems
	2018-06	iotedge	4	Cypress Semiconductor, Gigamon, Altera	20	Twilio, CA, Synopsys
	2018-07	microsoft-ui-xaml	3	Semtech, Diebold Nixdorf, Enernoc	16	Meta Platforms, Linkedin, Open Text
	2018-08	codeql	3	Pixelworks, Semtech, Stamps.com	20	Salesforce, Red Hat, Open Text
	2018-09	nni	3	Medidata Solutions, Liveramp Holdings, Zscaler	29	Red Hat, Altair Engineering, Intl Business Machines
	2018-09	Trill	5	Spansion, Liveramp Holdings, Mellanox Technologies	28	Meta Platforms, Open Text, Channeladvisor
	2018-10	winforms	1	Semtech	26	Box, 8x8, Meta Platforms
	2018-10	wpf	1	Liveramp Holdings	23	Salesforce, Citrix Systems, Open Text
	2018-12	onnxruntime	5	Mellanox Technologies, Spansion, Medidata Solutions	22	PTC, Intl Business Machines, Citrix Systems
	2018-12	accessibility-insights-windows	1	Semtech	12	Open Text, Box, Upland Software
	2019-01	calculator	7	Power Integrations, Smartsheet, Mellanox Technologies	18	Meta Platforms, Intuit, Verisign
	2019-03	Industrial-IoT	2	Semtech, Par Technology	17	Twilio, Open Text, Citrix Systems
	2019-03	microsoft-authentication-library-for-java	2	Smartsheet, Exlservice Holdings	19	Verisign, Earthlink Holdings, Meta Platforms
	2019-05	PowerToys	4	Appfolio, Smartsheet, GTY Technology Holdings	21	Intuit, Intl Business Machines, Open Text
	2019-11	BotFramework-Composer	2	Exlservice Holdings, Medidata Solutions	28	Box, Meta Platforms, Autodesk
	2019-11	runtime	1	Mellanox Technologies	33	Citrix Systems, CA, Salesforce
	2020-04	WindowsAppSDK	1	Liveramp Holdings	25	Meta Platforms, Open Text, Oracle
	2020-05	clarity	2	Medidata Solutions, Liveramp Holdings	27	Earthlink Holdings, Open Text, Linkedin
	2020-07	azurelinux	1	Stamps.com	31	Box, Ringcentral, Docusign
	2020-08	bicep	2	Cambium Learning Group, Liveramp Holdings	23	Box, Commvault Systems, Meta Platforms
	2020-09	onefuzz	3	Mellanox Technologies, Outerwall, Spansion	32	Open Text, Box, Commvault Systems
	2021-02	Extensible-Storage-Engine	2	Ping Identity Holding, Stamps.com	22	Intl Business Machines, Oracle, Box
	2021-03	venice	2	Liveramp Holdings, Enernoc	23	Intl Business Machines, Zendesk, Open Text

			Complement peers		Substitute peers	
Firm	Date	Repository	*N*	Highest on technology	*N*	Highest on technology
	2022-02	`go`	1	Liveramp Holdings	42	Verisign, Open Text, Box
	2022-04	`CleanReader.Desktop`	1	Pubmatic	–	–
	2022-05	`Microsoft-3D-Movie-Maker`	2	Srax, Smartsheet	15	Web.com Group, Verisign, Intl Business Machines
	2023-03	`data-api-builder`	3	GTY Technology Holdings, Dynatrace, Backblaze	31	Atlassian, Fastly, Meta Platforms
	2023-04	`AIShell`	9	Dynatrace, C3.ai, Bigbear.ai Holdings	41	Liveperson, Gen Digital, Servicenow
	2023-08	`DSC`	4	Gitlab, CS Disco, Dynatrace	27	Citrix Systems, Box, Red Hat
	2024-02	`sudo`	1	Gitlab	21	Atlassian, Meta Platforms, Intuit
	2024-03	`garnet`	9	Mellanox Technologies, Netlist, Ambarella	53	Intl Business Machines, Citrix Systems, Unisys
	2024-10	`openvmm`	2	Servicesource Intl, Dealertrack Technologies	21	Godaddy, Meta Platforms, Docusign
	2024-10	`BitNet`	1	Liveramp Holdings	22	NVIDIA, Altair Engineering, Intel
Mongodb	2018-11	`leafygreen-ui`	4	Light & Wonder, Semtech, Penguin Solutions	12	Adobe, Box, Microsoft
Morgan Stanley Dean Witter &	2017-06	`hobbes`	1	Radisys	28	Intl Business Machines, Red Hat, Salesforce
National Instruments	2017-05	`design-system`	2	Entropic Communications, NVE	13	Meta Platforms, Adobe, Autodesk
	2022-09	`nimble`	4	Light & Wonder, Smartsheet, Intralinks Holdings	11	Meta Platforms, Autodesk, Open Text
Netapp	2016-12	`trident`	3	Netlist, Nexttrip, Mitek Systems	26	Commvault Systems, Falconstor Software, Box
Netflix	2013-01	`ribbon`	3	Synopsys, Cantaloupe, Nuance Communications	8	Verisign, Open Text, Intl Business Machines
	2013-06	`genie`	2	Qlogic, Micros Systems	15	Open Text, Intl Business Machines, Citrix Systems
	2014-07	`security_monkey`	5	Micros Systems, Fair Isaac, Kratos Defense & Security	15	Red Hat, Intl Business Machines, Splunk
	2014-12	`atlas`	2	Intellicheck, Intralinks Holdings	7	Verisign, Intl Business Machines, Qumu
	2015-08	`falcor`	1	Identiv	21	Synacor, Verisign, Open Text
	2015-11	`spinnaker`	4	Infoblox, Corelogic, Kratos Defense & Security	17	Box, Open Text, Synchronoss Technologies
	2018-01	`iceberg`	7	Virtusa, Corelogic, Agilysys	20	Box, Intl Business Machines, Oracle
NVIDIA	2017-12	`aistore`	3	Enernoc, CSP, Corelogic	27	Commvault Systems, Box, Citrix Systems
	2018-06	`DALI`	2	Coherent, Alphabet	32	Advanced Micro Devices, Intl Business Machines, Intel
	2018-10	`server`	4	Rimini Street, Workiva, Super Micro Computer	29	Splunk, Citrix Systems, Red Hat
	2021-08	`NVFlare`	6	Jive Software, Widepoint, Lionbridge Technologies	33	Datadog, Rackspace Technology, Autodesk
	2021-08	`DLSS`	1	Pubmatic	4	Intl Business Machines, Snap, Alteryx
	2021-11	`NVIDIAImageScaling`	1	Exela Technologies	40	Datawatch, Intel, Advanced Micro Devices
	2022-01	`instant-ngp`	1	Intellicheck	44	Meta Platforms, Autodesk, Citrix Systems
	2023-03	`cuda-quantum`	4	Intellicheck, Smartsheet, Couchbase	33	Intl Business Machines, Advanced Micro Devices, Commvault Systems
	2023-10	`TensorRT-LLM`	1	Sprout Social	38	Intl Business Machines, Citrix Systems, Altair Engineering
Oracle	2013-01	`opengrok`	2	Semtech, Corelogic	9	PTC, Tigerlogic, Microsoft
	2014-09	`mysql-server`	5	Planet Payment, Semtech, Cambium Learning Group	18	Barracuda Networks, Edgio, Intl Business Machines
	2015-10	`mysql-workbench`	10	Semtech, Streamline Health Solutions, Light & Wonder	25	Box, Zendesk, Citrix Systems
	2016-01	`graal`	5	Quicklogic, Pixelworks, Maxim Integrated Products	32	Salesforce, Red Hat, PTC
	2016-06	`pgql-lang`	1	Widepoint	15	Box, Intuit, Intl Business Machines
	2017-06	`railcar`	5	Netlist, Liveramp Holdings, First Solar	28	Intl Business Machines, Red Hat, Open Text
	2018-09	`jdk`	2	Bankrate, NCI	19	Synacor, Box, Linkedin

			Complement peers		Substitute peers	
Firm	Date	Repository	*N*	Highest on technology	*N*	Highest on technology
	2019-04	jdk12u	1	NCI	13	Synacor, Edgio, Open Text
	2019-04	jdk11u	3	Smartsheet, Semtech, NCI	13	Synacor, Linkedin, Open Text
	2019-05	jfx	5	Liveramp Holdings, Medidata Solutions, Light & Wonder	24	Citrix Systems, Box, Open Text
	2019-10	labs-openjdk-11	3	Srax, NCI, Alphabet	16	Web.com Group, Progress Software, Open Text
	2019-12	jdk14	1	NCI	13	Synacor, Linkedin, Commvault Systems
	2020-06	coherence	23	Liveramp Holdings, Maxlinear, Altera	44	Domo, Box, Commvault Systems
	2020-09	labs-openjdk-15	3	Waitr Holdings, Ari Network Services, Bazaarvoice	10	Cadence Design Systems, Open Text, Intl Business Machines
	2020-11	nashorn	7	Cambium Learning Group, Liveone, Liveramp Holdings	36	Citrix Systems, Channeladvisor, Commvault Systems
	2020-12	jdk16u	2	Telaria, NCI	18	Synacor, Edgio, Linkedin
	2021-02	labs-openjdk-16	2	Waitr Holdings, Bazaarvoice	9	Jfrog, Cadence Design Systems, Open Text
	2021-02	jdk8u	3	Telaria, Rocket Fuel, NCI	16	Synacor, Meta Platforms, Linkedin
	2021-06	jdk17u	5	Pubmatic, Semtech, Smartsheet	14	Everbridge, Open Text, Synacor
	2021-09	jfx17u	3	Intellicheck, Spansion, Nanthealth	8	Mimecast, Synacor, Upland Software
	2021-12	jdk18u	2	Telaria, NCI	18	Synacor, Linkedin, Autodesk
	2022-06	jdk19u	3	Telaria, Rocket Fuel, NCI	22	Synacor, N-able, Autodesk
	2023-06	jdk21u	6	CS Disco, Smartsheet, Pubmatic	19	Atlassian, Edgio, Open Text
	2023-07	jfx21u	2	Intellicheck, Spansion	12	Atlassian, Edgio, Digitalocean Holdings
	2023-11	graalvm-community-jdk21u	12	Quicklogic, Credo Technology Group, Pubmatic	38	Atlassian, Cloudera, Commvault Systems
	2024-07	jfx23u	3	Intellicheck, Smartsheet, Appfolio	8	Atlassian, Edgio, Digitalocean Holdings
	2024-08	sandwood	11	Liveramp Holdings, Ambarella, Peloton Interactive	59	Intl Business Machines, Godaddy, Altair Engineering
Paypal Holdings	2023-04	junodb	4	Kratos Defense & Security, Cavium, KeyW Holding	28	Gen Digital, Elastic, Citrix Systems
Pier 1 Imports	2016-10	rocketbelt	8	Syntel, Mantech Intl, Infoblox	18	Box, Salesforce, Red Hat
Ping Identity Holding	2021-12	astro	7	Light & Wonder, Bentley Systems, Intralinks Holdings	22	Slack Technologies, Open Text, Microsoft
Red Hat	2018-09	porta	4	Exlservice Holdings, Pixelworks, Exela Technologies	22	PTC, Cloudera, Citrix Systems
Salesforce	2011-11	dataloader	2	Pixelworks, Electronics for Imaging	13	Gen Digital, Intl Business Machines, CA
	2015-09	design-system	2	Semtech, Super Micro Computer	26	Intl Business Machines, Citrix Systems, Synacor
	2017-10	design-system-react	5	Semtech, Global Payments, Block	24	Intl Business Machines, Adobe, Box
	2019-08	mule	19	Semtech, Mellanox Technologies, Analog Devices	51	Box, Verisign, Mobileiron
	2020-08	bolt-python	3	Electronics for Imaging, Penguin Solutions, Vonage Holdings	19	Meta Platforms, Open Text, Earthlink Holdings
	2021-09	pwa-kit	2	Semtech, Light & Wonder	26	Pinterest, Citrix Systems, Intuit
	2023-06	ampscript-core	10	Spansion, Texas Instruments, Backblaze	44	Citrix Systems, Atlassian, Edgio
	2024-10	NPSP	2	Semtech, Penguin Solutions	30	Blackbaud, Open Text, Workday
Schrodinger	2023-07	sketcher	7	Intralinks Holdings, Talend, Nutanix	37	Citrix Systems, Autodesk, Simulations Plus
Target	2019-11	server	4	Infoblox, Servicesource Intl, A10 Networks	16	Red Hat, Salesforce, Oracle
Teradata	2017-02	kylo	2	Code Rebel, Semtech	21	Commvault Systems, Box, Mobileiron
Twilio	2019-08	paste	2	Stamps.com, Enernoc	20	PTC, Salesforce, CA
Tyler	2022-08	forge	9	Cavium, Cadence Design Systems, Imperva	22	Dropbox, Autodesk, Synacor
Uber Technologies	2018-03	baseweb	4	Nuance Communications, Pareteum, KeyW Holding	10	Autoweb, Open Text, Web.com Group

			Complement peers		Substitute peers	
Firm	Date	Repository	*N*	Highest on technology	*N*	Highest on technology
	2023-05	`mock`	2	Coherent, Taskus	33	Atlassian, Box, Verisign
Workday	2019-07	`canvas-kit`	5	Semtech, Trubridge, Cavium	16	Blackbaud, Open Text, Synacor
Workiva	2019-05	`eva`	5	Semtech, Cavium, Penguin Solutions	31	Open Text, Intuit, Verisign

Table IA.4. External Validation of the Peer Classification

This table relates the peer classification to measures that are not used in building it. Listed as product-market rivals and listed as supplier or customer are from the Hoberg and Phillips (2016) TNIC and vertical relatedness data, in percent of pairs. Rivalry score is the TNIC similarity among listed pairs. Return correlations use daily returns over trading days −250 to −21 before the release: raw, in excess of the market, and Carhart (1997) four-factor residuals. Same four-digit SIC code is 100 for pairs with the same code and 0 otherwise. Correlation of annual sales growth uses the six fiscal years ending before the release. Pairs is the number of (release, peer) pairs. Panel A reports group means and differences in means, with standard errors clustered at the release level. Panel B regresses four of the measures on the technology and product-market percentile ranks with release fixed effects, with standard errors clustered at the peer-firm level. t-statistics are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Group Means of External Measures

	Pairs	Comp.	Sub.	Ctrl.	Comp. − sub.	Comp. − ctrl.
Listed as product-market rivals (per cent of pairs)	32,664	5.4	26.5	7.1	-21.1***	-1.7**
					(-13.50)	(-2.10)
Rivalry score, among the pairs that are listed	3,353	0.018	0.038	0.024	-0.020***	-0.007**
					(-7.34)	(-2.45)
Listed as supplier or customer (per cent of pairs)	32,664	2.0	1.4	1.1	+0.6	+0.9*
					(+0.99)	(+1.67)
Correlation of daily returns	24,855	0.276	0.368	0.248	-0.093***	+0.027***
					(-12.69)	(+3.62)
Correlation in excess of the market	24,855	0.024	0.079	0.023	-0.055***	+0.001
					(-8.35)	(+0.09)
Correlation of Carhart residuals	24,855	0.008	0.039	0.011	-0.031***	-0.003
					(-7.73)	(-0.82)
Same four-digit SIC code (per cent of pairs)	32,664	15.7	35.0	28.2	-19.3***	-12.5***
					(-5.89)	(-5.61)
Correlation of annual sales growth before the release	16,944	-0.033	0.123	0.043	-0.156***	-0.076***
					(-6.06)	(-3.08)

Panel B: External Measures Regressed on the Classification Axes

	Pairs	Technology rank	Product-market rank	Within R^2
Listed as a product-market rival (per cent)	32,664	+9.46	+23.84***	0.066
		(+1.63)	(+3.86)	
Listed as supplier or customer (per cent)	32,664	+0.97**	-1.05	0.000
		(+2.35)	(-1.03)	
Return correlation, excess of market	24,855	+0.0395***	+0.0675***	0.039
		(+3.40)	(+8.22)	
Return correlation, Carhart residual	24,855	+0.0069	+0.0395***	0.017
		(+1.22)	(+6.10)	

Table IA.5. Complement and Substitute Peers on Five External Measures

This table repeats Table 5 with release fixed effects only, so the comparison is between complement and substitute peers of the same release without holding the peer firm fixed, and adds two measures the body table omits. One measure is taken from each source. *TNIC rival* is 100 when the pair appears in the Hoberg and Phillips (2016) product-market competitor list and 0 otherwise. *Same SIC4* is 100 when the two firms report the same four-digit SIC code and 0 otherwise. *Sales growth co-movement* is the correlation of the two firms' annual sales growth over the six fiscal years ending the year before the release. *Return correlation* is the correlation of the two firms' daily Carhart (1997) four-factor residuals over trading days -250 to -21 before the release. *Vertical link* is 100 when the Hoberg and Phillips (2016) data link the two firms as supplier and customer and 0 otherwise, and is the one measure that should separate the groups the other way. Control peers are dropped, so the coefficient is the difference between a complement and a substitute peer and the two group means are printed below it. Standard errors are clustered on the peer firm; t-statistics are in parentheses; ***, **, and * denote significance at the 1%, 5%, and 10% levels.

Dependent variable =	TNIC rival (%)	Same SIC4 (%)	Sales growth co-movement	Return correlation	Vertical link (%)
	(1)	(2)	(3)	(4)	(5)
Complement	-18.85***	-21.63***	-0.1608***	-0.0315***	+1.05
	(-7.58)	(-8.53)	(-3.75)	(-6.06)	(+0.89)
Release FE	Yes	Yes	Yes	Yes	Yes
Peer-firm FE	No	No	No	No	No
Observations	6,457	6,457	4,277	5,380	6,457
Peer firms	425	425	295	389	425
Within R^2	0.035	0.045	0.012	0.014	0.001
Mean of dep. var., complement peers	5.39	15.67	-0.0329	0.0081	2.00
Mean of dep. var., substitute peers	26.49	35.00	0.1233	0.0389	1.36

Table IA.6.
Alternative Measures of Peer Firms' Sales

This table repeats the estimation of Table 7 with sales measured without scaling by assets: Panel A in millions of dollars, estimated by Poisson, and Panel B in logarithms. Columns (1) to (3) are those of Table 7, and so are the sample, the control peers' weights, the control variables, the fixed effects and the clustering. t-statistics are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)
Panel A: Sales ($m)			
Post × Complement	0.008		-0.037
	(0.35)		(-1.53)
Post × Substitute		-0.099***	-0.091***
		(-14.91)	(-14.01)
Complement − substitute			0.054**
			(2.23)
Wald p			0.025
Mean of dep. var.	1,503.724	11,804.228	1,503.724
Observations	91,651	112,939	116,971
Pseudo R^2	0.706	0.602	0.650
Panel B: Log sales			
Post × Complement	-0.0057		-0.0526***
	(-0.64)		(-5.90)
Post × Substitute		-0.0157***	-0.0054
		(-3.75)	(-1.40)
Complement − substitute			-0.0472***
			(-5.00)
Wald p			0.000
Mean of dep. var.	6.1557	7.4070	6.1557
Observations	91,577	112,866	116,897
Adjusted R^2	0.984	0.992	0.992
Controls	Yes	Yes	Yes
Cohort × Firm FE	Yes	Yes	Yes
Cohort × Year FE	Yes	Yes	Yes

Table IA.7.
Alternative Measures of Peer Firms' Profitability

This table repeats the estimation of Table 7 with profitability in place of sales, each over lagged total assets and one panel each. Columns (1) to (3) are those of Table 7, and so are the sample, the control peers' weights, the control variables, the fixed effects and the clustering. t-statistics are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)
Panel A: Gross profit			
Post × Complement	-0.0180***		-0.0136***
	(-3.90)		(-2.72)
Post × Substitute		0.0151***	0.0166***
		(5.04)	(6.17)
Complement − substitute			-0.0302***
			(-6.02)
Wald p			0.000
Mean of dep. var.	0.4368	0.4936	0.4368
Observations	91,631	112,917	116,949
Adjusted R^2	0.718	0.818	0.799
Panel B: Operating income			
Post × Complement	-0.0072**		-0.0115***
	(-2.36)		(-3.76)
Post × Substitute		0.0036**	0.0053***
		(2.09)	(3.43)
Complement − substitute			-0.0169***
			(-5.24)
Wald p			0.000
Mean of dep. var.	0.0642	0.0973	0.0642
Observations	91,607	112,892	116,924
Adjusted R^2	0.740	0.749	0.750
Panel C: ROA			
Post × Complement	0.0003		-0.0008
	(0.08)		(-0.19)
Post × Substitute		0.0034*	0.0039**
		(1.76)	(2.15)
Complement − substitute			-0.0046
			(-1.11)
Wald p			0.265
Mean of dep. var.	-0.0244	0.0147	-0.0244
Observations	91,631	112,917	116,949
Adjusted R^2	0.629	0.628	0.631
Controls	Yes	Yes	Yes
Cohort × Firm FE	Yes	Yes	Yes
Cohort × Year FE	Yes	Yes	Yes

Table IA.8.
Alternative Measures of Peer Firms' Investment and Employment

This table repeats the estimation of Tables 8 and 9 with investment and employment measured in other ways. Panel A is the total investment of Richardson (2006). Panels B and C are employment without sales, as a headcount estimated by Poisson and as its logarithm. Columns (1) to (3), the sample and the specification are those of Table 7. t-statistics are in parentheses; ***, **, and * denote significance at the 1%, 5%, and 10% levels.

	(1)	(2)	(3)
Panel A: Total investment			
Post × Complement	-0.0196***		-0.0141***
	(-3.86)		(-2.72)
Post × Substitute		-0.0019	-0.0016
		(-0.73)	(-0.64)
Complement − substitute			-0.0125**
			(-2.37)
Wald p			0.018
Mean of dep. var.	0.2067	0.2168	0.2067
Observations	91,631	112,917	116,949
Adjusted R^2	0.327	0.329	0.322
Panel B: Employment (thousands)			
Post × Complement	-0.003		-0.018
	(-0.19)		(-1.25)
Post × Substitute		-0.020***	0.002
		(-3.40)	(0.37)
Complement − substitute			-0.020
			(-1.47)
Wald p			0.141
Mean of dep. var.	5.151	28.756	5.151
Observations	91,281	112,563	116,586
Pseudo R^2	0.747	0.706	0.750
Panel C: Log employment			
Post × Complement	-0.0429***		-0.0688***
	(-5.39)		(-8.34)
Post × Substitute		0.0113***	0.0217***
		(2.94)	(5.95)
Complement − substitute			-0.0905***
			(-10.60)
Wald p			0.000
Mean of dep. var.	0.6203	1.6260	0.6203
Observations	91,281	112,563	116,586
Adjusted R^2	0.982	0.987	0.986
Controls	Yes	Yes	Yes
Cohort × Firm FE	Yes	Yes	Yes
Cohort × Year FE	Yes	Yes	Yes

Table IA.9. Alternative Measures of Peer Firms' Patenting

This table repeats Table 11 with patent applications. The dependent variable is the number of patent applications in Panel A and the number of applications classified as artificial intelligence in Panel B, dated by filing year. Applications are observed only if granted. Panel A covers filings through 2024 and Panel B through 2022. Coefficients are Poisson pseudo-maximum likelihood estimates. Column (1) compares complement peers and column (2) substitute peers with control peers, each excluding the other exposed group and with control peers balanced on that group alone; column (3) includes all three groups, balances control peers on complement and substitute peers together and estimates both coefficients jointly, and Complement − substitute is their difference with a Wald p-value. The sample, the control peers' weights, the control variables, the fixed effects and the clustering are those of Table 7. t-statistics are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)
Panel A: Patent applications			
Post × Complement	-0.015		-0.092**
	(-0.32)		(-2.34)
Post × Substitute		0.032*	0.019
		(1.69)	(1.06)
Complement − substitute			-0.111***
			(-2.90)
Wald p			0.004
Mean of dep. var.	24.485	386.553	24.485
Observations	89,951	110,844	114,815
Pseudo R^2	0.012	0.015	0.011
Panel B: AI patent applications			
Post × Complement	0.005		-0.068
	(0.08)		(-1.34)
Post × Substitute		-0.087***	-0.096***
		(-4.50)	(-5.19)
Complement − substitute			0.029
			(0.59)
Wald p			0.556
Mean of dep. var.	19.749	522.050	19.749
Observations	81,124	99,924	103,563
Pseudo R^2	0.027	0.071	0.065
Controls	Yes	Yes	Yes
Cohort × Firm FE	Yes	Yes	Yes
Cohort × Year FE	Yes	Yes	Yes

Table IA.10. Active and Passive GitHub Contributions of Peer Firms

This table repeats Table 12 with GitHub contributions split by type. The dependent variable is the number of active contributions (forks, issues, issue comments, pull requests and pushes) in Panel A and passive contributions (watch events) in Panel B, by the firm's accounts on repositories the firm does not own. Coefficients are Poisson pseudo-maximum likelihood estimates. Column (1) compares complement peers and column (2) substitute peers with control peers, each excluding the other exposed group and with control peers balanced on that group alone; column (3) includes all three groups, balances control peers on complement and substitute peers together and estimates both coefficients jointly, and Complement − substitute is their difference with a Wald p-value. The sample, the control peers' weights, the control variables, the fixed effects and the clustering are those of Table 7. t-statistics are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)
Panel A: GitHub active contributions			
Post × Complement	0.757**		0.634*
	(2.42)		(1.88)
Post × Substitute		0.188***	0.180***
		(3.62)	(3.56)
Complement − substitute			0.454
			(1.39)
Wald p			0.164
Mean of dep. var.	14.777	103.533	14.777
Observations	26,718	41,238	42,446
Pseudo R^2	0.045	0.010	0.010
Panel B: GitHub passive contributions			
Post × Complement	0.249***		0.330***
	(3.54)		(4.19)
Post × Substitute		0.314***	0.313***
		(14.56)	(14.97)
Complement − substitute			0.018
			(0.23)
Wald p			0.819
Mean of dep. var.	5.417	47.675	5.417
Observations	26,718	41,238	42,446
Pseudo R^2	0.020	0.036	0.039
Controls	Yes	Yes	Yes
Cohort × Firm FE	Yes	Yes	Yes
Cohort × Year FE	Yes	Yes	Yes

Table IA.11. Alternative Measures of Distress and Restructuring

This table repeats Table 13 with other measures of distress and restructuring. Special items, debt issuance and acquisitions are the Compustat items SPI, DLTIS and AQC, each scaled by lagged total assets. Altman Z is the Altman (1968) Z-score, with operating income after depreciation for earnings before interest and taxes; higher values mean lower default risk. Zmijewski PD is the normal transformation of the Zmijewski (1984) probit score. CHS PD is the logistic transformation of the Campbell et al. (2008) 12-month failure score, computed with annual inputs. Continuous scores are winsorized at 1% and 99% within fiscal year. The sample, the control peers' weights, the control variables, the fixed effects and the clustering are those of Table 7. t-statistics are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Dependent variable =	Special items	Debt issuance	Acquisitions	Altman Z	Zmijewski PD	CHS PD
	(1)	(2)	(3)	(4)	(5)	(6)
Post × Complement	0.0018	−0.0137*	−0.0045	0.629***	−0.0211***	0.0100*
	(1.23)	(−1.71)	(−0.98)	(4.00)	(−4.88)	(1.78)
Post × Substitute	−0.0010	−0.0012	−0.0003	0.395***	0.0094***	0.0158***
	(−1.54)	(−0.37)	(−0.12)	(6.67)	(4.73)	(4.83)
Complement − substitute	0.0028*	−0.0125	−0.0042	0.234	−0.0305***	−0.0057
	(1.85)	(−1.52)	(−0.89)	(1.49)	(−7.01)	(−0.95)
Wald p	0.064	0.130	0.373	0.137	0.000	0.344
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Cohort × Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Cohort × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep. var.	-0.0218	0.1455	0.0652	5.242	0.2127	0.0573
Observations	114,303	115,871	111,400	113,453	113,478	103,313
Adjusted R^2	0.132	0.237	0.163	0.742	0.855	0.350

Table IA.12. Placebo Labels

This table reports a randomization test of column (1) of Tables 7 to 12. In each draw the complement, substitute and control labels are reshuffled among the peers of each release, keeping the number of peers in each group, and Post $\times$ Complement is re-estimated on the reshuffled complement and control peers without weights. The observed row is the same regression on the true labels. Placebo mean and s.d. describe the distribution of the placebo coefficients, the rejection share is the fraction of draws with $|t| > 1.96$, and the placebo p-value is the fraction of draws whose $|t|$ is at least the observed $|t|$. Sales, investment (capital expenditure) and R&D are scaled by lagged total assets; Emp./sales is the number of employees per million dollars of sales; patents, AI patents and GitHub contributions are counts. Control variables are size, market-to-book and leverage. All regressions include cohort-by-firm and cohort-by-year fixed effects, where each release defines a cohort. Standard errors are clustered at the release-by-fiscal-year level. The sample covers the 249 releases from 2010 to 2024 and, for each release, fiscal years $t-3$ to $t+3$.

	Sales	Investment	Emp./sales	R&D	Patents	AI patents	GitHub
Observed Post $\times$ Complement	-0.0440^{***}	-0.0020^{***}	-0.2117^{***}	-0.0052^{***}	-0.215^{***}	-0.205^{***}	0.577^{***}
	(-5.33)	(-3.74)	(-3.43)	(-3.38)	(-4.76)	(-4.82)	(4.24)
Placebo mean	-0.0009	-0.0000	-0.0157	-0.0002	-0.002	-0.002	0.002
Placebo s.d.	0.0127	0.0008	0.1071	0.0021	0.052	0.052	0.099
Placebo $\lvert t\rvert > 1.96$, share	0.250	0.100	0.250	0.220	0.300	0.270	0.120
Placebo p-value	0.000	0.000	0.030	0.040	0.010	0.010	0.000
Draws	100	100	100	100	100	100	100

Table IA.13. Alternative Constants and Screens in the Peer Classification

This table reports the main estimates under alternative constants and screens of the peer classification. The first row of Panel A is the baseline. Each other row of Panel A moves one constant of the rule, and its last row removes the releasing firms from every peer pool. Each row of Panel B moves one screen. Peer pools hold 80 to 525 firms, so the nine-peer floor does not bind. Rel. is the number of releases with at least one complement and one control peer with return and accounting data. The first two columns are the value-weighted calendar-time abnormal returns of the complement portfolio, in percent, as in Table 6, with White standard errors. The remaining columns are Post × Complement for the outcomes of Tables 7 to 12, with the entropy-balancing weights solved again for every row. GitHub contributions are estimated by Poisson. The sample, the control variables, the fixed effects and the clustering are those of Table 7. t-statistics are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

		Complement portfolio		**Post × Complement**				
Classifier	**Rel.**	$[0,+1]$	$[0,+3]$	Sales	Investment	Emp./sales	R&D	GitHub
Panel A: The constants of the rule								
Reported classifier	249	-0.625^{***}	-0.734^{**}	-0.0602^{***}	-0.0017^{***}	-0.2410^{***}	-0.0050^{***}	0.5496^{**}
		(-3.34)	(-2.33)	(-7.00)	(-3.10)	(-5.25)	(-3.35)	(2.29)
Corners: top and bottom quarter	154	-0.492	0.340	-0.0888^{***}	-0.0042^{***}	-0.3640^{***}	-0.0115^{***}	1.1728^{***}
		(-1.43)	(0.66)	(-5.07)	(-3.80)	(-4.48)	(-4.22)	(3.42)
Corners: top and bottom 40%	293	-0.249^{*}	-0.291	-0.0306^{***}	-0.0016^{***}	-0.1436^{***}	-0.0024^{**}	0.4132^{***}
		(-1.86)	(-1.42)	(-5.42)	(-4.26)	(-5.31)	(-2.37)	(3.59)
Technology weights 0.50 / 0.50	264	-0.626^{**}	-0.722^{*}	-0.0545^{***}	-0.0026^{***}	-0.2029^{***}	-0.0065^{***}	0.5182^{**}
		(-2.44)	(-1.95)	(-6.37)	(-4.62)	(-4.48)	(-4.42)	(2.15)
Technology weights 0.80 / 0.20	240	-0.568^{***}	-0.512	-0.0514^{***}	-0.0015^{***}	-0.2103^{***}	-0.0041^{***}	0.5228^{**}
		(-2.86)	(-1.54)	(-6.03)	(-2.63)	(-4.65)	(-2.69)	(2.17)
Control technology at or below 40th pct.	249	-0.625^{***}	-0.734^{**}	-0.0569^{***}	-0.0016^{***}	-0.2426^{***}	-0.0039^{**}	0.4693^{**}
		(-3.34)	(-2.33)	(-6.42)	(-2.82)	(-4.99)	(-2.54)	(2.00)
Control technology at or below 60th pct.	249	-0.625^{***}	-0.734^{**}	-0.0569^{***}	-0.0018^{***}	-0.2407^{***}	-0.0050^{***}	0.4903^{*}
		(-3.34)	(-2.33)	(-6.67)	(-3.35)	(-5.37)	(-3.37)	(1.92)
Control product market at or below median	249	-0.625^{***}	-0.734^{**}	-0.0612^{***}	-0.0014^{**}	-0.2187^{***}	-0.0040^{***}	0.5516^{**}
		(-3.34)	(-2.33)	(-6.88)	(-2.38)	(-4.79)	(-2.66)	(2.25)
No vertical filter	250	-0.548^{***}	-0.473	-0.0607^{***}	-0.0019^{***}	-0.2294^{***}	-0.0049^{***}	0.5430^{**}
		(-2.90)	(-1.51)	(-7.01)	(-3.38)	(-5.22)	(-3.29)	(2.27)
Releasing firms removed from peer pools	257	-0.434^{**}	-0.318	-0.0515^{***}	-0.0012^{**}	-0.2502^{***}	-0.0032^{**}	0.3876^{***}
		(-2.14)	(-0.90)	(-5.79)	(-2.27)	(-6.52)	(-2.18)	(5.18)
Panel B: The screens and the peer pool								
Peer floor of 100 peers	241	-0.633^{***}	-0.745^{**}	-0.0606^{***}	-0.0018^{***}	-0.2339^{***}	-0.0050^{***}	0.5480^{**}
		(-3.36)	(-2.34)	(-6.98)	(-3.21)	(-4.85)	(-3.35)	(2.28)
At least three complement peers	190	-0.600^{***}	-0.560^{*}	-0.0634^{***}	-0.0018^{***}	-0.2219^{***}	-0.0057^{***}	0.5716^{**}
		(-3.26)	(-1.66)	(-6.95)	(-2.98)	(-4.29)	(-3.56)	(2.32)
Technology from the goods-text score alone	227	-0.557^{***}	-0.747^{***}	-0.0475^{***}	-0.0013^{**}	-0.2197^{***}	-0.0018	0.3824
		(-2.93)	(-2.71)	(-5.11)	(-2.22)	(-5.89)	(-1.18)	(1.54)
Control: every peer not in a corner	249	-0.625^{***}	-0.734^{**}	-0.0554^{***}	-0.0024^{***}	-0.2403^{***}	-0.0040^{***}	0.0194
		(-3.34)	(-2.33)	(-6.87)	(-4.68)	(-5.99)	(-2.74)	(0.07)
Peer pool without semiconductors	246	-0.557^{***}	-0.542^{*}	-0.0719^{***}	-0.0024^{***}	-0.0959^{**}	0.0007	0.3898^{*}
		(-2.62)	(-1.74)	(-7.10)	(-4.42)	(-2.19)	(0.54)	(1.92)

Table IA.14. Alternative Control Groups

This table repeats column (1) of Tables 7 to 12 with the control group weighted or matched in seven ways; complement peers always carry a weight of one. The reported row is the body specification, entropy balancing with complement peers as the target. The others are the common target of column (3), entropy balancing on the pre-release levels only, no weights, inverse propensity weights, coarsened exact matching on size, sales and investment strata, and propensity score matching of each complement peer to the three closest control peers of the same release. Imbalance is the largest absolute standardized difference across the balancing moments after weighting, and Eff. controls is the Kish effective sample size of the control group over its raw count. Patents, AI patents and GitHub contributions are estimated by Poisson. The sample, the control variables, the fixed effects and the clustering are those of Table 7. t-statistics are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	Balance		**Post × Complement**						
Control group	Imbalance	Eff. controls	Sales	Investment	Emp./sales	R&D	Patents	AI patents	GitHub
EB, complement target (reported)	0.000	0.79	-0.0381^{***}	-0.0031^{***}	-0.2438^{***}	-0.0037^{**}	-0.229^{***}	-0.212^{***}	0.554^{***}
			(-4.44)	(-5.80)	(-5.12)	(-2.47)	(-4.15)	(-4.36)	(2.58)
EB, complement and substitute target	0.716	0.28	-0.0616^{***}	-0.0019^{***}	-0.2431^{***}	-0.0057^{***}	-0.197^{***}	-0.192^{***}	0.550^{**}
			(-7.01)	(-3.30)	(-5.32)	(-3.82)	(-3.20)	(-3.79)	(2.29)
EB, levels only	0.027	0.79	-0.0406^{***}	-0.0031^{***}	-0.2382^{***}	-0.0034^{**}	-0.242^{***}	-0.224^{***}	0.551^{***}
			(-4.72)	(-5.89)	(-5.02)	(-2.32)	(-4.32)	(-4.60)	(2.58)
No weights	0.387	1.00	-0.0409^{***}	-0.0022^{***}	-0.3258^{***}	-0.0052^{***}	-0.217^{***}	-0.211^{***}	0.564^{***}
			(-4.78)	(-3.96)	(-6.29)	(-3.31)	(-4.67)	(-4.87)	(3.12)
IPW	0.010	0.78	-0.0249^{***}	-0.0028^{***}	-0.2518^{***}	-0.0030^{**}	-0.101^{**}	-0.082^{*}	0.380^{***}
			(-2.70)	(-5.24)	(-4.76)	(-2.02)	(-2.57)	(-1.82)	(3.13)
CEM	0.166	0.63	-0.0402^{***}	-0.0015^{***}	-0.2479^{***}	-0.0055^{***}	-0.198^{***}	-0.198^{***}	0.611^{***}
			(-4.66)	(-2.71)	(-5.25)	(-3.66)	(-4.10)	(-4.49)	(3.21)
PSM, 1:3	0.063	0.08	-0.0283^{***}	-0.0045^{***}	-0.2327^{***}	-0.0027	–	–	0.565^{***}
			(-2.59)	(-5.59)	(-4.04)	(-1.49)			(2.61)
Observations, reported			92,623	92,623	92,216	92,623	90,896	90,896	27,006

Table IA.15. Alternative Specifications and Samples

This table repeats column (1) of Tables 7 to 12, complement peers against entropy-balanced control peers, changing one element of the specification or the sample in each row. Clean control peers drops every control peer that is a complement or substitute peer of another release whose event year is within three years. Cluster on release clusters the standard errors at the release level and leaves the coefficients unchanged. Industry-by-year fixed effects adds fixed effects for the peer's two-digit SIC industry interacted with the fiscal year. The last three rows drop the releases of 2020 and 2021, the releases of 2023 and 2024, and the releases after 2019. The sample, the control variables, the fixed effects are those of Table 7. t-statistics are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Specification	Sales	Investment	Emp./sales	R&D	Patents	AI patents	GitHub
Baseline, column (1)	-0.0381^{***}	-0.0031^{***}	-0.2438^{***}	-0.0037^{**}	-0.229^{***}	-0.212^{***}	0.554^{***}
	(-4.44)	(-5.80)	(-5.12)	(-2.47)	(-4.15)	(-4.36)	(2.58)
Clean control peers	-0.0508^{***}	-0.0061^{***}	-0.0384	-0.0074^{***}	-0.517^{***}	-0.432^{***}	0.568^{**}
	(-5.48)	(-9.80)	(-0.70)	(-4.91)	(-7.46)	(-8.54)	(2.23)
Cluster on release	-0.0381^{***}	-0.0031^{***}	-0.2438^{***}	-0.0037	-0.229^{**}	-0.212^{***}	0.554^{**}
	(-3.36)	(-4.11)	(-3.39)	(-1.55)	(-2.36)	(-2.69)	(2.35)
Industry-by-year fixed effects	-0.0358^{***}	-0.0035^{***}	-0.1676^{***}	-0.0018	-0.074^{**}	-0.068^{**}	0.474^{**}
	(-4.20)	(-6.37)	(-3.55)	(-1.29)	(-2.01)	(-2.33)	(2.21)
Without 2020–2021 releases	-0.0424^{***}	-0.0034^{***}	-0.2991^{***}	-0.0073^{***}	-0.235^{***}	-0.225^{***}	0.727^{***}
	(-4.51)	(-5.72)	(-5.57)	(-4.45)	(-3.94)	(-4.39)	(2.65)
Without 2023–2024 releases	-0.0447^{***}	-0.0036^{***}	-0.2149^{***}	-0.0042^{**}	-0.238^{***}	-0.222^{***}	0.217^{**}
	(-4.73)	(-6.40)	(-4.06)	(-2.54)	(-4.23)	(-4.37)	(2.46)
Releases before 2020 only	-0.0575^{***}	-0.0044^{***}	-0.2611^{***}	-0.0094^{***}	-0.243^{***}	-0.234^{***}	0.446^{***}
	(-5.29)	(-6.75)	(-4.15)	(-4.92)	(-3.92)	(-4.38)	(3.61)
Observations, baseline	92,623	92,623	92,216	92,623	90,896	90,896	27,006

Table IA.16. Alternative Samples Excluding One Releasing Firm at a Time

This table reports the main estimates excluding one releasing firm at a time. Each row drops one releasing firm with at least five releases; the last row keeps only the releases of Microsoft Corp. Rel. is the number of releases in the subsample. The first two columns report value-weighted calendar-time abnormal returns of the complement portfolio, in percent, as in Table 6, with White standard errors. The remaining columns report Post × Complement for the outcomes of Tables 7, 8, 9, 10 and 12, from a regression on complement, substitute and entropy-balanced control peers, with the weights re-estimated in each subsample and standard errors clustered at the release-by-fiscal-year level. GitHub contributions are estimated by Poisson pseudo-maximum likelihood. Sales, investment (capital expenditure) and R&D are scaled by lagged total assets; Emp./sales is the number of employees per million dollars of sales; GitHub contributions are counts. t-statistics are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

		Complement portfolio		**Post × Complement**				
Sample	**Rel.**	$[0,+1]$	$[0,+3]$	Sales	Investment	Emp./sales	R&D	GitHub
All releases	249	-0.625***	-0.734**	-0.0603***	-0.0017***	-0.2410***	-0.0050***	+0.5496**
		(-3.34)	(-2.33)	(-7.00)	(-3.10)	(-5.25)	(-3.35)	(+2.29)
Without Microsoft Corp	179	-0.600***	-0.672*	-0.0567***	-0.0011*	-0.2037***	-0.0053***	+0.4589*
		(-2.59)	(-1.73)	(-5.73)	(-1.67)	(-3.69)	(-3.05)	(+1.69)
Without Alphabet Inc	209	-0.802***	-0.994***	-0.0629***	-0.0016**	-0.1881***	-0.0030*	+0.4345*
		(-3.90)	(-2.89)	(-6.67)	(-2.57)	(-3.97)	(-1.92)	(+1.69)
Without Oracle Corp	222	-0.553***	-0.621**	-0.0700***	-0.0019***	-0.2391***	-0.0060***	+0.5594**
		(-2.89)	(-1.99)	(-7.62)	(-3.33)	(-4.87)	(-3.80)	(+2.20)
Without Meta Platforms Inc	230	-0.587***	-0.590*	-0.0539***	-0.0018***	-0.2607***	-0.0050***	+0.4939**
		(-3.09)	(-1.83)	(-6.28)	(-3.10)	(-5.45)	(-3.25)	(+2.06)
Without Amazon.com Inc	234	-0.708***	-0.848***	-0.0618***	-0.0017***	-0.2462***	-0.0061***	+0.5811**
		(-3.76)	(-2.65)	(-6.91)	(-2.80)	(-4.99)	(-3.87)	(+2.24)
Without Nvidia Corp	240	-0.541***	-0.644**	-0.0590***	-0.0017***	-0.2418***	-0.0053***	+0.5532**
		(-2.86)	(-2.01)	(-6.83)	(-2.93)	(-5.14)	(-3.48)	(+2.26)
Without Salesforce Inc	241	-0.607***	-0.706**	-0.0614***	-0.0018***	-0.2555***	-0.0052***	+0.5404**
		(-3.18)	(-2.20)	(-6.93)	(-3.23)	(-5.36)	(-3.37)	(+2.14)
Without Netflix Inc	242	-0.614***	-0.718**	-0.0615***	-0.0019***	-0.2514***	-0.0051***	+0.5458**
		(-3.23)	(-2.25)	(-6.99)	(-3.24)	(-5.41)	(-3.34)	(+2.26)
Without International Business Machines	243	-0.593***	-0.665**	-0.0597***	-0.0019***	-0.2462***	-0.0055***	+0.5488**
		(-3.17)	(-2.11)	(-6.83)	(-3.34)	(-5.27)	(-3.63)	(+2.29)
Without Intel Corp	243	-0.598***	-0.826***	-0.0614***	-0.0020***	-0.2446***	-0.0048***	+0.5466**
		(-3.18)	(-2.80)	(-7.01)	(-3.49)	(-5.23)	(-3.14)	(+2.28)
Microsoft Corp alone	70	-0.485	-0.589	-0.0756***	-0.0038***	-0.3810***	-0.0045	+1.0028***
		(-1.57)	(-1.46)	(-4.40)	(-3.37)	(-5.22)	(-1.62)	(+3.48)

B Appendix Figures

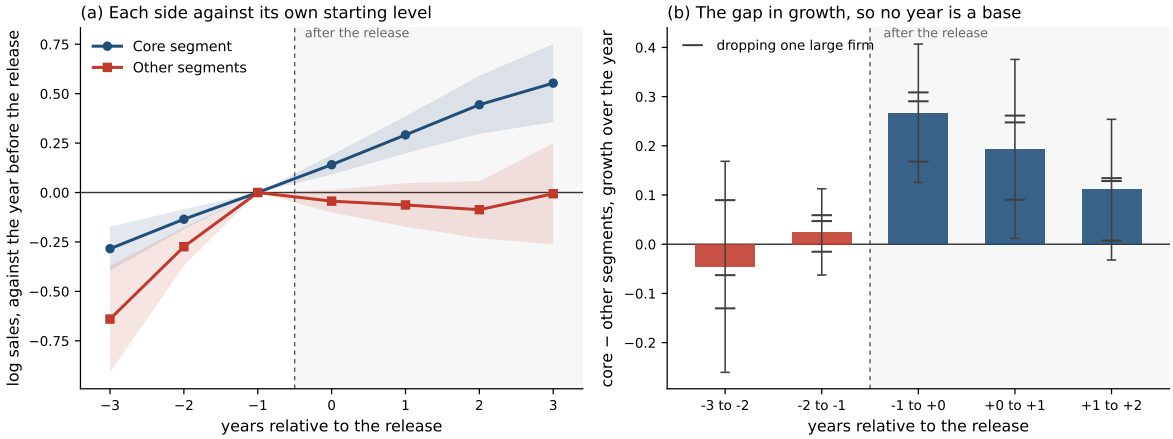

Figure IA.1. The Releasing Firm's Own Segments Around the Release
This figure plots the Compustat business-segment sales of releasing firms around releases. The core segment is the segment with the largest average sales over the three fiscal years before the release; all other segments are aggregated. Only segments reported in year $t-1$ are included. Panel (a) plots the logarithm of sales of each side relative to year $t-1$. Panel (b) plots the difference in annual sales growth between the two sides from a regression with release-by-side and release-by-year fixed effects. Bars are 90% confidence intervals based on standard errors clustered at the releasing-firm level. Markers report the same coefficient excluding each of the three firms with the most releases.

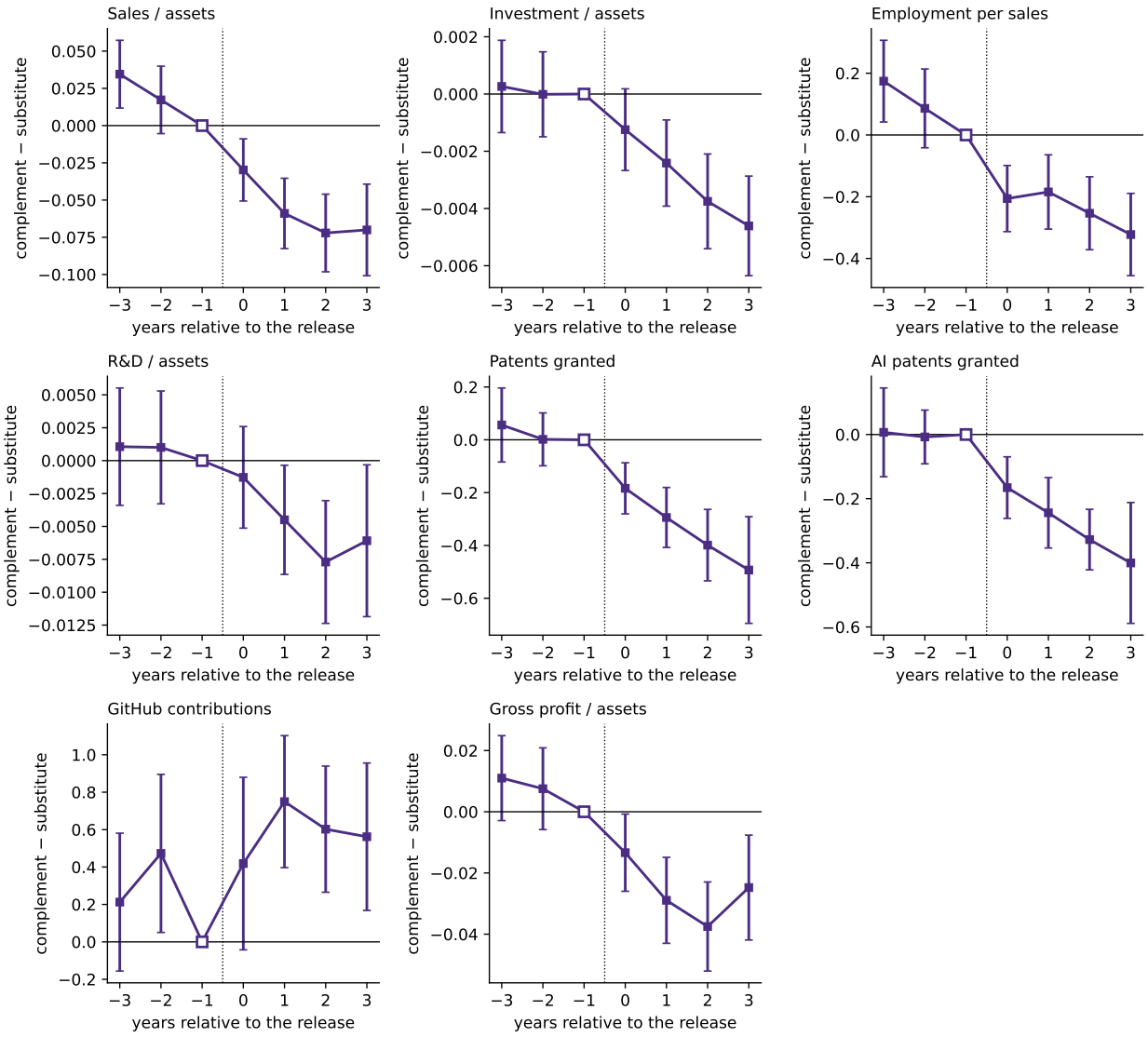

Figure IA.2. Complement Against Substitute Peers, Every Outcome

This figure plots event-study coefficients comparing complement peers with substitute peers, estimated as in column (3) of the corresponding body tables. Each panel reports the coefficient on the interaction of an event-year indicator with the complement indicator, with control peers excluded. Patents and GitHub contributions are estimated by Poisson pseudo-maximum likelihood. Year $t-1$ is the omitted category. Control variables are size, market-to-book and leverage. All regressions include cohort-by-firm and cohort-by-year fixed effects, where each release defines a cohort. Bars are 90% confidence intervals based on standard errors clustered at the release-by-fiscal-year level.