Igniting Innovation: Evidence from PyTorch on Technology Control in Open Collaboration*

Daniel Yue
Georgia Institute of Technology

Frank Nagle
Harvard Business School

This draft: April 2025

Abstract. Many companies offer free access to their technology to encourage outside add-on innovation, hoping to later profit by raising prices or harnessing the power of the crowd while continuing to steer the direction of innovation. They can achieve this balance by opening access to the technology (access rights) but still maintaining governing control over it (control rights). However, how this continued exertion of control influences other companies' choice to invest in furthering that technology is not well understood. This study looks at the impact of technology control on external contributions in open collaboration contexts by examining the case of PyTorch, a popular machine learning framework, which shifted its governance from a for-profit corporation (Meta) to a non-profit foundation in 2022. The results show that this shift led to a small decrease in contributions from Meta but a notable increase from external companies. In particular, participation increased from complementors (Chip Manufacturers); by contrast, users (App Developers and Cloud Providers who rely on PyTorch as input) did not change their rate of participation. These findings are consistent with the notion that the governance change resolved complementors' hold-up concerns.

*Correspondence to: dyue@gatech.edu. We thank Shane Greenstein, Carliss Baldwin, Becky Karp, Christian Peukert, and Jeff Fossett for their helpful comments, as well as seminar participants at the HBS D^3 Institute Research Day and the Open User Innovation Conference. We also thank Linux Foundation employees Jen Shelby, Kylie Wagar-Dirks, and Lucy Hyde for providing additional context on the PyTorch governance change. Professor Nagle's research lab receives funding from both the Linux Foundation and GitHub. However, neither organization had a role in the design, analysis, or editorial decisions related to this manuscript. All errors remain our own.

1 Introduction

Open collaboration has long been studied as a model for distributed innovation, inspired by examples like Wikipedia or Linux, (Baldwin and von Hippel 2011; Levine and Prietula 2014; Ren et al. 2016; Kane and Ransbotham 2016; Gambardella and Von Hippel 2019). However, the past decade has seen a rise in the strategic use of open collaboration in innovation efforts by firms across prominent settings like electric vehicles, digital platforms, and artificial intelligence[1], raising new economic questions. In particular, *firm-sponsored* open collaboration is characterized by the fact that, although the focal firm makes its technology open for use and follow-on innovation by the public, it retains various forms of control over the technology. Such firm-sponsored projects are paradoxically open (in the sense of free access for use and modification) but not *open* (in the sense of decentralized technical control). Due to the possibility of coordinated, strategic action on the part of the focal firm, it is likely that community participation, particularly external firm participation, will vary based on the strength of the focal firm's control rights. How do technology control rights impact an external firm's likelihood of participation in open collaboration? Which firms are more sensitive to technology control rights?

Understanding the relationship between control rights and external firm participation in firm-sponsored open collaboration is essential because of the breadth of relevant economic settings where it occurs. For example, firms have been shown to strategically open technology (such as committing to not enforce intellectual property rights) to increase control over technology standards, improving competitive positioning in supplier networks (Augereau et al. 2006; Jones et al. 2021; Toh and Pyun 2024). Many companies sponsor the development of open source software (OSS) for key technologies to encourage development and interoperability with related software (O'Mahony and Karp 2020; Haese and Peukert 2024; Azoulay et al. 2024), including recent developments in Generative Artificial Intelligence (GenAI) models. Finally, digital platforms may be seen as open collaborations when they leverage governance decisions like data sharing or low initial royalty rates to attract complementors and increase complementor compatibility with the platform (Wareham et al. 2014; Rietveld et al. 2019).

However, despite its potential economic importance, the effect of technology control rights on firm participation in open collaboration is understudied in the prior literature. Preliminary but limited evidence (Boudreau 2010; O'Mahony and Karp 2020) suggests that a focal firm's reliquishing of control rights does impacts an external firm's likelihood of participating in open collaboration. However, the mechanism through which this occurs is debated[2]. Relatedly, to our knowledge, there is no direct evidence on how this effect may differ depending on the type of external firm, specifically, whether they are a user of the open technology, or a complementor to it, a key dichotomy in inter-firm relationships.

This literature gap likely exists because of three key challenges in empirically estimating this relationship. The first and most subtle challenge is that changes in "control rights" almost always correlate with changes in what the literature calls "access rights" — the direct ability for third-party firms to appropriate value freely or at low cost[3]. The prior literature

[1]Examples are discussed throughout the paper, but top of mind examples include Tesla's 2014 release of its electric vehicle battery technology (Musk 2014), Google's sponsorship of the cloud containerization software Kubernetes (Evans 2018), or Meta's sponsorship of PyTorch, the focal setting of this paper (Meta 2022).

[2]We describe our paper's relationship and contribution to this literature at length below.

[3]As a prominent example, Tesla's opening up of its electric vehicle patents reduced Tesla's control over

(Boudreau 2010; Schwarz and Takhteyev 2010) mostly studies this distinction in the context of platforms that grant access to complementors to build on the core technology, equating control with the open-sourcing of the core platform (and therefore the inevitable opening of access rights). Only recent research by O'Mahony and Karp (2020) demonstrates that governance structures of already open-sourced technologies represent a further aspect of technology control, one that can vary independently of access rights. However, finding such settings that cleanly separate the influence of these two related effects in this way is difficult. Second, in many contexts, it can be difficult to find sufficient statistical power to study differences in the effect of technology control rights on different external participants. Finally, the third challenge is that it is hard to find plausibly exogenous variation in control rights to convincingly identify a causal effect, as shifts in control rights may be anticipated or even endogenously determined by community actions (e.g., through the lobbying of government regulators).

This paper addresses these challenges by studying an economically important open source project (i.e., with constant access rights) as it changed its governance model from firm-sponsored ("dominant") governance to community-led ("distributed") governance. Specifically, we study PyTorch, a leading machine learning framework, as it unexpectedly transitioned governance from the for-profit corporation Meta to the Linux Foundation (a non-profit organization focused on supporting open source software (OSS) projects) in September 2022. PyTorch's governance change presents a unique opportunity to test the effect of technology control rights on external firm participation because PyTorch was *already* open source at the time of the change and had been so for almost its entire existence dating back to 2016. Therefore, this governance shift did not change the direct ability of external firms to appropriate value from the technology, but *did* change their ability to control the future trajectory of the technology. We interpret this governance change as a shift in control rights but not access rights, allowing us to separate the two effects and solve the first challenge highlighted above. Additionally, there is a large amount of external firm contributors, allowing us to address the second challenge. Motivated by our theoretical framework, we study differences in participation between external firms that are users versus those that are primarily complementors. Finally, PyTorch's announcement was sudden and unanticipated by the machine learning community. When applicable, we leverage this plausible "exogenous to the external community" property to justify causal interpretations of our estimated effects and make progress on the third challenge.

We operationalize this empirical design by assembling a novel, user-level panel data set from GitHub (the primary website hosting OSS projects) of all historical contributions to PyTorch. Descriptive analysis focuses on characterizing overall contribution trends to PyTorch around the time of the governance change. Our data comprises 4,138 unique contributors (33.9% from Meta[4]) to PyTorch from April 2020 - September 2023. Notably, contributors include a long tail of relatively low-commitment individuals, which we use as a helpful comparison group in some parts of our analysis. Instead, we focus primarily on employee contributors from the largest and most recognizable external firms contributing to PyTorch (other than Meta), including NVIDIA, Hugging Face, Apple, and Microsoft. We identify 964 total contributors employed at 72 key external firms responsible for 7.19% of our dataset's total commits[5]. These data allow us to estimate difference-in-difference

the technology but also directly gave access to the technology to other car manufacturers (Musk 2014).

[4]Meta's contributors are more active than other contributors on average; Meta contributions account for over 70% of the total commits to PyTorch during the timeframe of the study.

[5]A commit is a group of simultaneous code changes made to a software project leveraging the Git version

specifications to test our key hypotheses.

Our empirical results highlight that control rights have an immediate, large, and sustained effect on firm contributions to PyTorch. The first group of results is from within-PyTorch difference-in-difference models estimated on variation over time in contributor-level participation. They show that the transition of PyTorch governance from Meta to the Linux Foundation leads to only a small increase in net contributors to the project. The small net effect conceals important heterogeneity: external firm participation rose by approximately 25% and remained elevated, while Meta's participation initially increased before dropping sharply after the governance change. This result highlights that shifting control rights to collective governance may not always increase total welfare because such a shift likely reduces the focal company's incentive to contribute while simultaneously increasing the incentives of other companies to contribute.

The second set of results highlights that the governance shift led to a sustained contribution increase for only a certain type of external firm. Namely, contributions increased from Chip Manufacturers, who began or increased making technology-specific investments needed to create interoperability between PyTorch and their computer chips. However, there were no changes in contributions by Non-Chip Manufacturers ("Application Developer" or "User") firms, who contribute primarily to learn and to improve the usability of the technology for themselves. We interpret these results as evidence that PyTorch's shift in control rights mitigated an existing hold-up problem for external firms. However, because "Complementor" Chip Manufacturers depend on continued interoperability of their chips with the software to capture value, they are more susceptible to hold-up by Meta and therefore experience a relative increase in participation (compared to "Users") when the threat is mitigated by the governance change. This result is robust to a triple-difference comparison, where we leverage TensorFlow (another popular open source machine learning framework controlled by Google) contributors as a control group.[6]

Our findings primarily contribute to the strategy literature on value capture in open innovation (Teece 1986; Chesbrough and Bogers 2014; Tambe 2014; Alexy et al. 2018; Nagle 2019; Rotolo et al. 2022), by being the first (to our knowledge) to apply a control rights lens to analyze firm participation in that setting. The closest of these papers to ours is Alexy et al. (2018), but it focuses on the perspective of the firm choosing to open their technology and why they might do so. We take this decision as given and consider how external firms are likely to react to this decision. More broadly, whereas the prior literature emphasizes the role of *direct* value appropriation in predicting firm participation in open collaboration (often via control of complementary assets), our paper is the first to empirically demonstrate that *future* value appropriation (through ex-post control rights) drive present-day firm participation decisions in open collaboration innovation systems.

Moreover, we contribute to the literature on firm participation in the governance of the digital commons (Ostrom 1990; West and O'Mahony 2008; Boudreau 2010; He et al. 2020; O'Mahony and Karp 2020; Altman et al. 2022; Tang et al. 2023). Our paper is most similar to two of these. Boudreau (2010) studies participation by hardware complementors in an operating system as a function of 'openness' and 'control', but identifies control with public access / lack of IP rights. By contrast, our paper studies a setting where the technology is entirely open source (i.e., in the public domain), but where control is determined

control system. For those without prior exposure to Git, you can think of a commit as a set of code changes tagged by the contributor and date contributed.

[6] For reasons detailed in the manuscript, we only use TensorFlow in triple diff (and not diff-in-diff) comparisons due to the presence of pre-trends in the event studies corresponding to diff-in-diff specifications.

by governance structure. Consequently, our main finding (relinquishing control increases participation) differs from the main finding of Boudreau 2010. O'Mahony and Karp (2020) study how changes in governance of IBM's Eclipse IDE affected external firm participation. Their explanation of external firm reactions to a platform opening technology focuses on concerns around appropriation by IBM, whereas we focus on hold-up due to control rights. More broadly, this literature largely emphasizes that the more 'open' the commons, the better. Our control rights lens adds nuance to this work by showing that shifting control rights to a more 'distributed' model does not necessarily lead to an increase in contributions as it may dampen the incentive for focal company investments while increasing the incentives to contribute for complementors, and having a limited impact on user contributions.

Lastly, our results contribute to the literature on firm participation in OSS communities (West and Lakhani 2008; Dahlander and Magnusson 2008; Nagle 2018; O'Mahony and Karp 2020; Murciano-Goroff et al. 2021; Nagaraj and Piezunka 2024; Fleischmann et al. 2023; Haese and Peukert 2024; Kim et al. 2024), by being the first paper (to our knowledge) to study how external firms responses to actions by open collaborations directly sponsored by other firms differ depending on whether they are complementors or users. **Note, we could add a bit more to distinguish from Haese/Peukert and Nagaraj/Piezunka, but I think this already sums it up sufficiently.***

We proceed as follows. In section 2, we develop a conceptual framework for analyzing the effect of control rights on firm participation in open collaboration. In section 3, we introduce PyTorch as the empirical domain for our tests, as well as our data and empirical design. In Section 4, we present the main results of our paper. And in section 5, we discuss several implications of our findings and conclude.

2 Conceptualizing Technology Control and Open Collaboration

In this section, we review the literature on control rights and firm participation in firm-sponsored open collaboration to theorize hypotheses to take to the data. We first define open collaboration, then define control rights in the context of open collaboration, and finally theorize their effect on external participation, especially that of external firms.

2.1 Defining Control Rights in Open Collaboration

Open collaboration refers to an important model of innovation where an open technology is developed by a distributed community, often without explicit contracts or research agreements (Baldwin and von Hippel 2011; Levine and Prietula 2014). Open collaboration differs from more traditional innovation models like research agreements or joint ventures because it requires many participants (not just two or three) to realize the full value-creation potential. Our interest is particularly in *firm-sponsored* open collaboration, where a focal firm releases a technology as a public good in order to encourage follow-on innovation by other firms. This variant is further distinctive because innovation is asymmetric between the focal firm and the follow-on firms. Such collaborations are especially common due to digitization, which brought about communication technologies that enabled the decentralized coordination of larger technology projects (Gambardella and Von Hippel 2019). Most prominently, both firm-led software and hardware development have occurred under an open-source model (e.g., at Meta, see Lin 2021 and https://opensource.fb.com/). However, other examples include companies intentionally releasing IP to encourage follow-on innovation (and possibly standardization around the technology), such as in the case of DuPont (Murray et al. 2016), Celera (Williams 2013), or Tesla (Musk 2014). In the same spirit (though perhaps not fully 'open' in the full sense of the word), platforms may provide low prices and open source software development kits (SDKs) to encourage complementary innovation, such as Hugging Face's model hub or various 'App' stores on platforms like Shopify.

What is the primary predictor of external firm participation in firm-sponsored open collaboration? Currently, the literature on open collaboration emphasizes *current* value appropriation channels ("access rights"). (Boudreau 2010; O'Mahony and Karp 2020). Access rights may include low royalty rates to benefit from a platform or use of intellectual property or the ability to use a software or hardware design in follow-on innovation for free. But what about expectations about *future* value capture (which we call "control rights"), where a firm may make a costly effort in order to provide optionality later on after markets develop?

In economic theory, a standard framework for reasoning about uncertain future value falls under a set of ideas known as *property rights theory* (Grossman and Hart 1986; Hart and Moore 1990).[7] In particular, property rights theory (also known as incomplete contracts theory) highlights that although many economic collaborations can be made efficient through the re-allocation of value through contracting (including those with quantifiable future uncertainty), in special economic situations of interest, relationship-specific effort and investment ("ex-ante investment") must precede the resolution of *non-contractible* uncertainty. In such situations, a critical governance decision is the allocation of *control rights* — the ex-ante ability (e.g., they are allocated before firms invest any effort) to make ex-post decisions (e.g., after non-contractible uncertainty is resolved) about how shared assets are used. The theory predicts that firms without ex-post decision rights will underinvest in ex-ante effort (relative to the social optimum) because of fear of the potential of "hold-up" by the firm with the control rights. The theory predicts that hold-up threats are stronger when those without ex-post decision rights have weaker outside options (i.e., value captured in the absence of ex-post cooperation). In this framework, control rights are optimally allocated when they are controlled by the firm whose investment has the greatest effect on net value creation. We thus adopt a definition of control rights as the ability for a single agent (often a single firm) to make decisions about an innovative asset *in the future* after current, non-contractible uncertainty is resolved.

In practice, where do control rights come from? While originally referring to ownership of physical property and used to study the boundary of the firm, the innovation literature has applied control rights theory to settings like bilateral research agreements and joint ventures (Aghion and Tirole 1994; Oxley 1997; Sampson 2004; Lerner and Malmendier 2010; Rodríguez and Nieto 2016) where the relevant control rights are decision rights about the R&D project (e.g. project termination) rather than explicit property rights, and where relationship-specific investments are interpreted as *technology-specific* investments. In that setting, control rights theory helped to identify optimal contract design for research collaborations that properly incentivized cooperation between firms (Lerner and Malmendier

[7] We utilize the term control rights and the associated theoretical framework, but our ideas also relate to a sibling literature in economics known as *Transaction Cost Economics*. In that literature, the focal governance structure is ex-post governance and monitoring rather than ex-ante decision rights. While the ideas are similar and, in some cases, overlapping, we prefer the control rights framework because of its sharp predictions around hold-up effects, which are the focus of this manuscript.

2010).

Our theoretical contribution is to apply incomplete contracts theory to a new innovation setting: open collaboration. Incomplete contracts theory is particularly applicable to open collaboration because there is often non-contractible uncertainty over how to capture value. This is because technical capabilities are rapidly shifting, meaning that business models are changing, regulations are forming, and other clear market structures have yet to take shape (Gao and McDonald 2022). Because of this, it's common for firms to prioritize value capture optionality while they wait for the regulatory and business model structure to solidify (McDonald and Gao 2019). Furthermore, many instances of open collaboration preclude the ability to formally contract on expected outcomes due to community norms (e.g., in the example of open source software).

In the context of open collaboration, control rights are most clearly allocated based on formal governance rights — explicit rules about who has decision-making power. However, control rights can also be developed through more informal social authority, such as through past technical contributions or expertise. For example, Python's creator, Guido Van Rossum, had control over the (open source) programming language's technical trajectory due to his authority as the creator, a role that was only later semi-formalized into his title as the "Benevolent Dictator for Life" (BDFL) (Van Rossum 2008). More generally, He et al. (2020) study the process by which OSS project licenses are changed and find that they are driven through discussions by core contributors to each project, often through "reflective agency" that focuses on emphasizing shared values across the project contributors.

2.2 Theorizing the Effect of Control Rights on External Firm Participation

The literature often describes the allocation of control rights as centralized or distributed (O'Mahony and Karp 2020). Just because a firm is sponsoring a technology does not mean the control rights are centralized; it is possible that control rights are seen as more distributed if the focal organization is seen as reputable, fair, cooperative, or ideologically motivated, such as the open source machine learning company Hugging Face (Greenstein et al. 2023). What is the effect of an exogenous shift in control rights from a centralized to a distributed model (in particular, holding present value-capture channels constant) on external firm participation?

The key insight of our conceptual framework is that "opening" control rights in this way does not create incentives out of thin air but rather shifts incentives between parties. In firm-sponsored open collaboration, the *focal firm* is often responsible for creating the raw infrastructure and rules through which other firms engage. By contrast, *external firms* are those that build on the focal innovation to create additional value. For example, Fontana and Greenstein (2021) demonstrate that the establishment of the Centrino Wifi standard by Intel (the focal firm) led to an influx of complementary products by router manufacturers (the external firms). Assuming that there is any expected value in these decision rights, such a change will, therefore, differentially impact these two groups.

Hypothesis 1. *In open collaboration, a change in governance from firm-dominant to collective will increase participation by external firms.*

Notably, external firms are different from (external) *unaffiliated* individuals, a class of contributors that have been extensively studied in the literature (West and Lakhani 2008; Dahlander and Magnusson 2008). In particular, Hypothesis 1 predicts the opposite of the literature on individual participation in open source ecosystems, which highlights that participants engage in open source development due to a mix of extrinsic preferences like the ability to become known and hired by the focal company or intrinsic preference of problem solving and skill-building enabled by the focal company's coordinating role (Lerner and Tirole 2002; Lakhani and Wolf 2005; David and Shapiro 2008; Shah and Nagle 2020; Tang et al. 2023). Hypothesis 1 also resonates with the core finding of O'Mahony and Karp (2020), which highlights that clear, collective governance led to deepened participation by external firms in the context of the open sourcing of IBM's Eclipse platform. We extend beyond such existing work in our consideration of the types of external firms below.

Although the focus of this paper is on the impact on such a governance shift on external firms, we can also consider the impact on the focal firm itself. Although the governance shift is certainly not exogenous to them, it is still interesting to consider how they are likely to act after such a decision. Such a consideration highlights an unappreciated downside of collective governance: it reduces the incentives for a focal firm to invest as they now have less at stake in the future of the project. Therefore, the effect of such a governance change on *net* contributions to a project is actually ambiguous. If a focal firm is a disproportionate contributor to a focal technology, then a shift in governance rights may actually reduce welfare by decreasing the focal firms' incentives to continue investing in the technology. This is equivalent to a "no free lunch" argument, because it shows that openness does not magically create incentives for external participation without associated costs, but rather shifts incentives between focal and external firms. In particular, control rights theory emphasizes that consideration of the optimal allocation of control rights (with respect to overall welfare) depends on the marginal returns to the ex-ante effort of each party (Grossman and Hart 1986; Hart and Moore 1990).

2.3 Heterogeneous Firms Motives for Participating in Open Collaboration

We distinguish between two types of external firms based on how they capture value from their contributions. The first class of external firms is those who derive value from follow-on innovation through the creation and control of profitable complementary products (per traditional arguments by Teece (1986) and its descendants, e.g., Alexy et al. (2018)). We simply call them *Complementors*. By contrast, a second class of external firms is those who contribute in order to increase the value of the technology to their own applications they are building on top of the focal technology. In particular, it has been shown that firms that contribute to open technologies can learn to use the technology in a more productive manner (Nagle 2018). We call such firms *Users*. User firms, therefore, get value from contributing as long as they can continue to access the technology[8].

While both types of firms make technology-specific investments (and are thus susceptible to hold-up problems), the key difference between Complementors and Users is the strength of their outside options (and the corresponding magnitude of hold-up). Recall

[8]We use the term *Complementor* and *User* to align with Porter's classic strategy framework (Porter 1980) where users or customers utilize a firm's product directly but complementors are firms whose products enhance the value of the focal firm's product when a user uses both products together. However, readers familiar with the platform literature may note that contributions from *both* types of external firms are presumably complementary (in terms of value creation) to the focal technology, adding confusion to the label "Complementor" (which can apply to both groups). Such readers can substitute the concepts of Complementor and User for the more recent concepts of Horizontal and Vertical complementors (respectively) in the platform literature, e.g., Thomas et al. (2024).

that in property rights theory, an outside option is the value an agent captures if ex-post cooperation between parties breaks down. Crucially, property rights theory predicts that when an agent has a stronger outside option, the magnitude of a hold-up problem is lessened. In open collaboration, such a breakdown in ex-post cooperation corresponds most closely to disagreement on the technical direction of a project. In that case, Users have a stronger outside option because they only depend on the interface (or "API") of the technology, shared across all users. Therefore, even if the focal firm and the User do not cooperate after investments are made, Users can still capture substantial value from their investment because their access rights are guaranteed (since the technology is open) and the technology's interface is unlikely to change significantly. By contrast, Complementors have a weaker outside option because they technically depend on specific integrations that enable the coupling between the focal technology and their complementary products to create value.[9] If the focal firm and Complementors do not cooperate after initial investments, the focal firm may limit or undo interoperability with the Complementor's product, meaning Complementors could not capture value from their investment and, therefore, experience greater hold-up. Formally stated:

Hypothesis 2. *The positive effect of a change in governance from firm-dominant to collective on external firm participation in open collaboration will be greater for Complementors than Users.*

Hypothesis 2 is critical for our empirical tests that interpret a governance change as a shift in control rights because it is singularly predicted by a control rights interpretation of that change. While a governance change may have several possible causal pathways that impact an external firm's involvement, we expect to see this pattern only if control rights are the primary mechanism underlying the effect.

3 Empirical Setting, Design, and Data

In this section, we explain PyTorch's relevance as an empirical setting for testing our hypotheses, describe our data collection process, and introduce our estimation strategy.

3.1 PyTorch as an Empirical Setting

3.1.1 Background on Open Source Machine Learning and PyTorch

To test our hypotheses, we need an empirical setting that features a technology developed by a focal firm, contributed to by multiple external firms, and containing variation in "control rights" that holds "access rights" constant. While challenging to find in general, a particularly useful setting where these properties hold is that of open source software, and in particular open source machine learning (OSML). Open source software is useful as a setting because it is by definition freely available, meaning that access rights to the technology are fixed. In particular, OSML is useful as a setting because its development is dominated by leading technology firms (for example, see Baruffaldi and Poege (2024)) and

[9] Essentially, we are arguing here that the nature of a firm's technical dependence (Baldwin and Clark 2000) predicts the strength of that firm's outside option. When a firm is only dependent on the general functionality of a technology, they have a stronger outside option; when a firm is dependent on a specific technical integration with the technology, they have a weaker outside option.

it features a framework (PyTorch) that experienced a sudden shift in control rights, as we explain below.

What is PyTorch? PyTorch is a machine learning framework used to specify, train, and run inference on neural networks[10]. In particular, PyTorch provides two key features that differentiate it from standard numerical computing libraries like Numpy or Matlab: 1) flexible specification and automatic differentiation of model architecture, enabling application to many modalities of data (numeric, text, speech, video, etc.) and 2) optimized training on different hardware accelerators (e.g. GPUs) and computing environments (e.g. distributed compute clusters). In other words, PyTorch makes it straightforward to operate neural networks at scale, a technique that has become increasingly popular given recent research breakthroughs leveraging models with billions of parameters.

PyTorch has arguably become the *most* popular machine learning framework, providing a technical backbone for a majority of artificial intelligence (AI) research and applications today. Paperswithcode.com reports that almost 70% of machine learning research papers with code use PyTorch (see Figure A1). All major cloud providers provide customized pre-packaged environments for leveraging PyTorch (AWS, Azure, GCP). Beyond direct usage, PyTorch also underlies many implementations of popular AI models today as both a critical dependency of higher-level open source libraries like Hugging Face's Transformers and Lightning AI's PyTorch Lightning, as well as direct usage by a diverse range of startups and larger companies including OpenAI, Elasticsearch, Tesla, Airbnb, Genentech, and Disney (Meta 2022, OpenAI 2020, PyTorch Foundation 2023, Sharma 2024, Lahoti 2019). While other machine learning frameworks exist and have similar functionalities (most notably TensorFlow, a Google-backed open source machine learning framework)[11], PyTorch won "market share" due its superior usability[12] and its strong technology ecosystem (He 2019). Undoubtedly, PyTorch has been essential to the emergence of AI technologies as a driver of the economy.

Critical to our purposes, PyTorch is open source under the Berkeley Software Distribution (BSD-3) license, where "redistribution and use in source and binary forms, with or without modification, are permitted" (see the source code), as long as appropriate attribution is given to the original creators in any forks (derivatives) of the code. The main repository's source code is publicly available on GitHub at `pytorch/pytorch`, and the broader `pytorch/` GitHub organization hosts several other repositories that cover a variety of complementary developer activities to the main repository (such as documentation, technical integrations, testing infrastructure, etc.).

[10] Understanding of machine learning terminology and PyTorch's functionality is not necessary for understanding this paper. However, for the interested reader, relevant definitions can be found in many places on the internet, including Google's Machine Learning Glossary.

[11] Indeed, PyTorch is the technical descendant of two such frameworks that came from academia: Lua Torch and Caffe2. There are many other similar software packages capable of implementing neural networks (and other popular machine learning techniques), such as Matlab, scikit-learn, OpenNN, Julia, Apache's MXNet, or Microsoft's CNTK (see Langenkamp and Yue 2022 for background). However, none of them have achieved the same level of community engagement and economic relevance as Meta's PyTorch and Google's TensorFlow, with the potential exception of Google's JAX.

[12] As a famous tweet by Andrej Karpathy in 2017 goes, "I've been using PyTorch a few months now and I've never felt better. I have more energy. My skin is clearer. My eye sight has improved."

3.1.2 PyTorch's Institutional History and Governance Change

Institutionally, PyTorch was conceptualized and developed by Meta (née Facebook) engineers (see Soumith's blogpost or Spisak's keynote).[13] It was the second major corporate-backed open source machine learning framework to emerge, with Meta's public release of PyTorch in 2017 following Google's public release of TensorFlow in 2015. PyTorch was developed as a community-driven project from its very beginnings, with many contributions coming from outside of the Meta organization, including from other companies like NVIDIA and Quansight.[14] Nevertheless, Meta has consistently provided the majority of investment in PyTorch: Table 1 shows that Meta employees have provided over 70% of commits to the repository since its 2017 public release.[15]

Table 1. Affiliation Breakdown of PyTorch Contributions. This table summarizes contributions to the PyTorch repositories by affiliation type—Meta, External Company, or Other—across the entire history of the project (until 2024). Some percentages do not sum to 100% due to bot-related contributions, which are not shown here.

Affiliation	Total Commits	% of Commits	Unique Committers	Total Issues	% of Issues	Unique Issue Authors
Meta	88001	70.98	1942	96433	52.77	1191
External Company	22221	17.92	734	27206	14.89	721
Other	13759	11.10	3640	57895	31.68	20949

Due to Meta's special relationship with PyTorch, prior to 2022, the company controlled the project at both the technical and strategic levels. PyTorch has tight technical integration with Meta's infrastructure – Meta famously sources its own production PyTorch code directly from the head of PyTorch's main branch, meaning that the GitHub PyTorch repository is technically a mirror of the "true" Facebook-internal repo, and thus only current Meta employees can "land" pull-requests (i.e., approve changes to the codebase that others pull from — see Eric Yang's blog or PyTorch documentation).[16]. Beyond technical control, Meta management had complete governing control over the project's technical vision and direction. This type of control is best understood as "soft power", in the sense that PyTorch prides itself on a clear division between its business and technical leadership of the project. Indeed, almost all public descriptions of the governance process highlight its democratic nature and emphasize technical, data-driven arguments over consideration of particular users or use-cases (Chinatala 2022).

In September 2022, Meta shocked the machine learning world by announcing a transition

[13] More generally, Meta is known for its broad support of open source development. For example, Meta is the primary firm sponsor of several other well-known open source projects, such as the React javascript library for developing responsive web interfaces, the GraphQL API language (link), several popular pre-trained machine learning models including Detectron2 and the Llama LLM family, and the ELI5 Open Compute project for developing open source designs for datacenter hardware (Lin 2021).

[14] Quansight is a data science and engineering company founded by Travis Oliphant, the creator of Numpy and the Anaconda organization.

[15] We discuss how we impute institutional affiliations of contributors later in Section 3.2.1. This amount of Meta contribution is a lower bound because some commits are tagged with generic email hosts like Gmail or Outlook, who may actually be Meta employees. Nevertheless, OSS norms would generally imply that individuals contributing on behalf of their employer use their work email address, validating this measurement approach.

[16] PyTorch leaders like Eric Yang and public documentation come off as genuinely apologetic about this arrangement, but we note that this arrangement has not changed as of 2024/07 despite at least three years of recorded public recognition of this problem. For reference in the documentation, see the answer to "Q: Can I become a committer on the project?"

of PyTorch's governance model to a standalone PyTorch Foundation, operated under the guidance of the Linux Foundation (LF). This governance change was made with the explicit intention to attract investment (both code and financial contributions) from external organizations into PyTorch "to accelerate progress in AI research" (Chintala 2022). The Linux Foundation, a non-profit organization established in 2000 to support Linux development and OSS projects, has a mission focused on ensuring neutrality across organizations on technical projects, including the popular operating system Linux and container orchestration system Kubernetes.

The hallmark of such LF-run projects is the presence of a governing board that votes to democratically make high-level decisions about both business and technical strategy (including who else to admit to the board). Thus, concurrent to the announcement of the PyTorch foundation, the organization brought on NVIDIA, AWS, Google Cloud, Microsoft Azure, and AMD as part of the shared governance board (Linux Foudation 2022, Meta 2022), a board that would be expanded throughout 2023 to include IBM, Hugging Face, Intel, Graphcore, Lightning AI, Huawei, and Snowflake (PyTorch 2023). All announcement posts emphasize that the board focuses on decisions related to business governance, not day-to-day technical governance of the tool, which was intended to remain unchanged.[17] Nevertheless, by changing the governance to a model run by a voting board of other organizations and bringing in the LF, Meta's singular control of the technical direction of the project (and potentially its social status as the creator of the tool) was greatly diluted.

The exact timing of when the 'effects' of the governance change begin is challenging to precisely pin down. Conversations with PyTorch Foundation insiders revealed that this transition had been discussed privately between Meta and LF for over two years, with the agreement being put in place in principle in Spring 2022 and a handful of other key organizations being recruited to the foundation around that time. Because we do not have a record of which companies were approached as potential members of the PyTorch foundation in Spring 2022, we expect that PyTorch's potential transition to the Linux Foundation was 'in the air' for these early joiners as early as Spring 2022, and structure our subsequent analysis accordingly.

Conceptually, we interpret this governance change as a shift in control rights from Meta to many external companies, what the literature calls a change from "dominant" to "collective" governance (O'Mahony and Karp 2020). Because collective governance requires voting and alignment of preferences between companies (similar to the Internet Engineering Task Force standard setting committees of Simcoe (2012)), Meta no longer has the ability to single-handedly set the technical direction of the tool (e.g., with respect to system architecture and higher level technology strategy). Therefore, they no longer have the ability to "hold up" a contributing company in the event that Meta's commercial interest diverges from the contributing company's interest. For example, in the event of a technical incompatibility, the PyTorch governance board may weigh in on whether the project should maintain support for specific back-end compilers; previously, Meta could have single-handedly decided to discontinue support for certain complementary technologies.

[17] Beyond the governance board, in practice, the LF primarily provides support in community organization through marketing, conference management, and administrative coordination. Technical leadership remains with the original maintainers of the tool.

3.1.3 Meta's Strategy and Competitive Considerations

Why did Meta invest so heavily in PyTorch in the first place, and why did they decide to spin the project out as a jointly governed foundation under the Linux Foundation in 2022? From the perspective of developing empirical tests for the effect of technology control rights, Meta's rationale for these transitions is orthogonal to this study (save for testing Hypothesis 2) as long as the change was unanticipated by external firms and unaffected by external firm strategizing. Nevertheless, Meta's technology strategy is a first-order managerial question, and we provide a more extended (speculative) discussion of this in Section A.1. Germane to the core theme of this paper, Meta leadership has explicitly called out control of the technology's future as a rationale for their original investment (PyTorch 2023). Therefore, the most thematic interpretation of Meta's decision to spin out PyTorch in 2022 is that sufficient non-contractible uncertainty had been resolved as time passed from its founding in 2017 to the governance transition in 2022. Under this rationale, as value capture channels from AI technologies became increasingly clear and contractible, Meta's expected marginal benefit of investment with singular governing control has diminished to the point where relinquishing governing control to encourage external firm investment became relatively attractive versus retaining technological control.

An astute reader may speculate that the governance change may have changed the competitive logic of contributing to PyTorch, differentially affected companies in direct competition with Meta (e.g., in social media or digital advertising). A priori, this seems plausible, and may well occur in other settings where governance of this type changes occur. However, while it seems like this empirical setting may support a competition-based interpretation, the results actually show that the governance change did not induce cooperation from competitors, but rather incentivized entry from complementors and users of the technology. We know this because only two companies that contribute to PyTorch over the time period studied could remotely be characterized as in direct competition with Meta. The first, Google, is a technology giant that competes with Meta in digital advertisement, but its contributions to PyTorch were purely limited to creating technical compatibility with its customized chip for machine learning ("Tensor Processing Units"); in this sense, Google is actually acting as a complementor here. ByteDance (the parent company of TikTok) competes with Meta over social media users, but its contribution magnitude is relatively small (8 total commits over the analysis period for this paper). Furthermore, technology competition is notoriously challenging to measure, and no clear consensus has emerged from the literature for how to do this. For this reason, our paper focuses on theorizing external firms as complementors to, or users of, PyTorch, and does not theorize competitive considerations.

3.2 Data and Measures

One of the distinct advantages of studying an open source software project like PyTorch is that a full contribution history is stored in a public git repository, available on GitHub.[18]

[18]Git is a source control management (SCM) system that allows software projects with multiple developers to track changes over time and ensure that the program continues to function as expected. Git (or similar SCM systems) are almost universally used and considered best practice for software projects due to the fact that code systems grow tremendously complex over time and small changes can easily break the overall system. Germane to our purposes, Git tracks granular information on contributors and the contributions that they make to the project over time. While the exact mechanics of git are somewhat complicated and can be found in detail online (see here), previously unexposed readers can think of commit changes as "code

By leveraging the public GitHub API, we gather data on the PyTorch organization's entire contribution history[19], including technical contributions data and contributor-level data. We aggregate the raw commit-level data[20] to the contributor-month level[21] to form our primary analysis data set. For our analysis, we focus on the timeframe between April 2020 and September 2023, which includes two years before the start of the transition period and one year after the public announcement of the governance change. In total this gives us 3.5 years (42 months) of data, covering the time surrounding the event of interest. As a control, we gather analogous data from TensorFlow, Google's open source machine learning framework.

3.2.1 Imputing Institutional Affiliations

Testing our hypotheses requires measuring company affiliation at the contributor level. To do this, we leverage the fact that the git version control system requires contributors to provide an associated "author" email address. We use this email in two ways. First, GitHub indirectly uses these emails to link commits to the author's corresponding GitHub profile. Because a significant benefit to open source contributions is the public recognition given to key contributors (and potential career benefits), these profiles tend to be populated and well maintained – allowing us to gather meta-data at the contributor level. In particular, contributors tend to self-report their affiliations in their profiles. The second benefit of the email data is that we can use the email domain to directly assign a company affiliation to some subset of the committers. There is a very strong norm in OSS development that contributors should use their work email address if they are contributing on behalf of their employer and a personal email address if they are contributing on their own time. For example, PyTorch founder Soumith Chintala's early contributions to PyTorch were made using the author's email "soumith@fb.com," his Facebook work email address, since creating PyTorch was his official job at Facebook.

Therefore, we impute company affiliations at the contributor level using the following steps. First, we gather self-reported affiliations from contributors' profiles and extract known institution names ("entity resolution") from these descriptions. Second, we take the domain of any email addresses associated with that contributor on GitHub via commits (e.g., "fb.com" or "nvidia.com") and develop a mapping between those domains and known companies. Lastly, for contributors with greater than 100 commits, we manually inspect the GitHub and LinkedIn profiles of contributors and record their company affiliations. We do this to ensure accurate affiliation information for the most active contributors. We then merge this information and record the contributor's affiliation as 1) Meta if any source reports the contributor as a Meta employee at any point, 2) any company if that contributor was identified as an employee at any point, 3) any university if the contributor was associated

change" level data associated with a software repository, tagged by the contributor and date contributed.

[19]PyTorch's organization on GitHub comprises 73 repositories, including the primary tool `pytorch/pytorch` and a constellation of related projects. Because of the high level of integration between PyTorch organization projects and overlap in the contributors, we view these projects as collectively comprising "PyTorch". However, as shown in the appendix, results are robust to analysis of only `pytorch/pytorch`.

[20]A commit is a code change by an author at a point in time, created for version control. While we gather data on commits, pull requests, issues, and comments, our analysis focuses on commits (in line with the literature, e.g., Wright et al. (2023)). However, our results are largely the same for using pull requests, with pull requests simply being a more aggregated version of commits. Issues and comments are more indicative of using the technology than contributing to it, so we end up not using that data in this analysis because our core interest is in firm contributions to the technology, rather than usage.

[21]Results are robust to analysis at the quarterly level.

with a particular school, and otherwise 4) unknown, a common outcome for accounts only associated with generic email domains like gmail.com or outlook.com. Importantly, our aggregation rolls up subsidiaries to the parent organization level: for example, we assign employees of Oculus or Instagram to Meta and DeepMind or Google Brain to Google.

Our approach is biased towards over-reporting contributors as Meta employees, which (if anything) would downwardly bias the results that we find later on (we exclude Meta employees from our test of Hypothesis 1 and **??**). The approach minimizes false-positive matches, leaving open the possibility of false-negatives (unaffiliated contributors who are actually employed), who contribute due to company influence but go unmarked due to the fact that they contribute using an unaffiliated email address (e.g., Gmail). The additional information from manual labeling is essential for our analysis due to the need for accurate affiliation-level data on contributions. We demonstrate the benefit of this approach for our analysis by comparing it to the literature's standard approach of just using email domains (for example Wright et al. (2024)) in Table A1. In particular, we show that different corporate affiliation implies a different likelihood of using a corporate email (versus self-reporting that affiliation in a GitHub or LinkedIn profile but using one's personal email to sign commits), highlighting the importance of the additional information captured by our method. As a robustness check, we later demonstrate that our main results are strengthened if we impute affiliation based only on the email domain, indicating that the results of our broader approach err on the side of underestimating the true effect.

3.2.2 Which Companies Contribute to PyTorch and TensorFlow?

A list of the top 25 companies (by total unique contributors) contributing to PyTorch or TensorFlow[22] is presented in Table 2. Notably, the top contributing companies comprise technologies companies, especially Chip Manufacturers (Intel, NVIDIA, ARM, AMD, IBM, etc.), Cloud Providers (Microsoft, Amazon, Alibaba, etc.), as well as other App Developers (Red Hat, OpenAI, Hugging Face, Spotify, etc.). Contributions to PyTorch and TensorFlow are dominated by contributions from Meta and Google (respectively), with more contributors from those organizations than any other organization.

Motivated by our theoretical framework in section 2, we classify the companies contributing to PyTorch as Chip Manufacturers[23] or Non-Chip Manufacturers. We interpret Chip Manufacturers as Complementors to PyTorch because their commercial incentive is to increase demand for their focal chips through improving the complementary software, a classic example of 'commercializing the compliment" (e.g., Teece (1986)). By contrast, Non-Chip Manufacturers can be either Application Developers (who use the technology in their own products) or Cloud Providers (who provide use of the technology on their computers) are interpreted as Users[24], getting their value primarily from learning the tool and increasing its applicability within their own products.

[22] We include TensorFlow here and later on in our analysis as a control group. TensorFlow is often seen as the competitor ML framework to PyTorch; this table shows that in practice, many companies contribute technically to both frameworks.

[23] Chip Manufacturers are companies whose primary focus is the design and production of semiconductor components, including CPUs, GPUs, and other integrated circuits, which serve as the critical processing units in computers and other electronic devices.

[24] We use the term Non-Chip Manufacturers throughout the manuscript for technical accuracy, but the reader can substitute the term 'Application Developer' if they find it a more clear description of the types of firms contained in this category. Later in the analysis, we provide a data-driven justification for grouping App Developers and Cloud Providers

Table 2. Top 25 Companies (by total unique contributors) contributing to PyTorch and TensorFlow from April 2020 - September 2023 (the analysis period covered in this paper). Note that PyTorch and TensorFlow refer to all repositories under the `pytorch/` and `tensorflow/` GitHub organizations, not just the primary repositories (`pytorch/pytorch` and `tensorflow/tensorflow`). The column 'Both' denotes the amount of contributors that commit to both PyTorch and TensorFlow in the analysis period. Not all PyTorch Board members joined on the first announcement; the full list and dates joined are listed here, as derived from the PyTorch Foundation's website. Note that Huawei (App Developer, October 2023), Snowflake (Cloud, December 2023), and Rebellions (Chip Manufacturer November 2024) also joined the PyTorch board at some point, but were not in the top companies contributing here.

			Contributors				Commits		
Company	PyTorch Board	Company Type	PyTorch	TensorFlow	Overlap	Total	PyTorch	TensorFlow	Total
Meta	Sep 2022	Focal Firm	1402	10	2	1410	52851	247	53098
Intel	Oct 2023	Chip Manufacturer	83	85	4	164	1058	2280	3338
Microsoft	Sep 2022	Cloud	80	21	1	100	1473	177	1650
NVIDIA	Sep 2022	Chip Manufacturer	65	47	2	110	4028	2153	6181
Amazon	Sep 2022	Cloud	47	14	4	57	1081	149	1230
Google	Sep 2022	Chip Manufacturer	44	1395	13	1426	1614	83478	85092
AMD	Sep 2022	Chip Manufacturer	29	24	1	52	695	724	1419
IBM	Jul 2023	Chip Manufacturer	23	19	1	41	171	174	345
Quansight		App Developer	22	0	0	22	3703	0	3703
Yandex		Cloud	11	6	1	16	17	19	36
Alibaba		Cloud	10	10	1	19	21	50	71
Fujitsu		App Developer	9	1	0	10	76	4	80
Hugging Face	Aug 2023	App Developer	9	3	0	12	12	71	83
Graphcore	Sep 2023	Chip Manufacturer	8	7	0	15	24	39	63
Tencent		App Developer	7	11	2	16	31	131	162
Apple		Chip Manufacturer	6	5	1	10	236	77	313
ARM	Sep 2024	Chip Manufacturer	6	50	3	53	14	1021	1035
ByteDance		App Developer	6	13	2	17	8	110	118
Cerebras		Chip Manufacturer	6	4	1	9	33	10	43
Lightning	Oct 2023	App Developer	6	0	0	6	107	0	107
Cruise		App Developer	5	0	0	5	186	0	186
GitHub		App Developer	5	3	0	8	24	19	43
OpenAI		App Developer	5	7	1	11	41	209	250
Preferred Networks		App Developer	5	0	0	5	53	0	53
Uber		App Developer	5	2	1	6	15	3	18

The key prediction of our theoretical framework is that Chip Manufacturers are more concerned about being 'held up' by Meta and, therefore, are most likely to contribute more after control rights are dispersed via the governance change. The reason is that Chip Manufacturers must maintain interoperability between their chips and PyTorch's implementation, whereas Non-Chip Manufacturers are only dependent on the PyTorch library interface. As a result, Chip Manufacturers are more concerned about loss of interoperability (and therefore value capture) should Meta be non-cooperative after investments are sunk. Indeed, discussions with the Linux Foundation highlighted that support for various chip backends was indeed a type of key concern that was brought to the governing board for strategic discussion, consistent with this prediction. Section 4.2 generalizes and tests for this insight in the data.

Table 3. Summary Statistics. All observations are at the Contributor-Month level. Panel A describes all contributors to PyTorch organization repositories. Panel B describes Non-Meta Contributors to PyTorch, the analysis sample used to test Hypothesis 1 and Hypothesis 2. Panel C describes a more limited sample: External Company contributors to PyTorch and TensorFlow; this is used as a more stringent test of **??**.

Variable	Mean	Stdev	Min	Max	N	# Contributors
Panel A. Sample: All Contributors to PyTorch						
Repo = PyTorch	1	0	1	1	173796	
Month			2020-04-01	2023-09-01	173796	
Post	0.2857	0.4518	0	1	173796	
Transition	0.1429	0.3499	0	1	173796	
Contributor					173796	4138
Meta	0.3388	0.4733	0	1	173796	1402
Unaffiliated	0.4886	0.4999	0	1	173796	2022
University	0.035	0.1839	0	1	173796	145
External Company	0.1375	0.3444	0	1	173796	569
Chip Manufacturer	0.064	0.2448	0	1	173796	265
Non-Chip Manufacturer	0.0735	0.2609	0	1	173796	304
1(Is Active)	0.0809	0.2727	0	1	173796	
Log(Commits + 1)	0.1129	0.4498	0	4.9053	173796	
Panel B. Sample: Non-Meta Contributors to PyTorch						
Repo = PyTorch	1	0	1	1	114912	
Month			2020-04-01	2023-09-01	114912	
Post	0.2857	0.4518	0	1	114912	
Transition	0.1429	0.3499	0	1	114912	
Contributor					114912	2736
Meta	0	0	0	0	114912	
Unaffiliated	0.739	0.4392	0	1	114912	2022
University	0.053	0.224	0	1	114912	145
External Company	0.208	0.4059	0	1	114912	569
Chip Manufacturer	0.0969	0.2958	0	1	114912	265
Non-Chip Manufacturer	0.1111	0.3143	0	1	114912	304
1(Is Active)	0.0573	0.2324	0	1	114912	
Log(Commits + 1)	0.0669	0.3197	0	4.804	114912	
Panel C. Sample: External Company Contributors to PyTorch and TensorFlow						
Repo = PyTorch	0.5602	0.4964	0	1	40488	
Month			2020-04-01	2023-09-01	40488	
Post	0.2857	0.4518	0	1	40488	
Transition	0.1429	0.3499	0	1	40488	
Contributor					40488	964
Meta	0	0	0	0	40488	
Unaffiliated	0	0	0	0	40488	
University	0	0	0	0	40488	
External Company	1	0	1	1	40488	964
Chip Manufacturer	0.5	0.5	0	1	40488	482
Non-Chip Manufacturer	0.5	0.5	0	1	40488	482
1(Is Active)	0.1272	0.3332	0	1	40488	
Log(Commits + 1)	0.1748	0.5337	0	4.804	40488	

3.2.3 Measures

For our analysis, we form a balanced panel of Contributor-Months for contributions to PyTorch organization repositories, summarized in Table 3, Panel A. Here, we describe the variables found within at greater length.

Time Variables (Transition and Post). In addition to using months as fixed effects, we form variables denoting two key periods of interest: the "Transition" period (April 2022–September 2022) and the "Post" period (On or After October 2022). We call the first period the "Transition" period because, as described in Section 3.1.2, conversations with the Linux Foundation emphasize that a handful of external contributors were looped into private discussions of the potential governance change during these months. Second, we consider the "Post" period as after October 2022 after the governance change based on the timing of the public announcement of the change.

Contributor Variables (Affiliation Labels). Across PyTorch, we find 4,138 unique contributors hailing from 802 unique institutional affiliations (as measured following the procedure described in Section 3.2.1). From these affiliation labels, we split contributors into four mutually-exclusive dummy variables: Meta employees, Unaffiliated (e.g. if contributed from an gmail.com address), University affiliates, or External Company employees. For External Company employees, we further form (mutually exclusive) dummy variables denoting whether the company is a Chip Manufacturer or not. Notably, there are *many* Unaffiliated contributors to PyTorch (2,022 in total), but this should not be confused with the idea that unaffiliated contributors drive PyTorch development. Most unaffiliated contributors commit once and then leave, whereas most Meta and External Company contributors are consistent contributors. Finally, we note that of the External Company contributors, Chip Manufacturer employees comprise about half of the contributing population.

Dependent Variables. We measure two outcome variables. First, 1(Is Active) measures whether a contributor committed within a given month, e.g., the extensive margin of participation. Second, Log(Commits + 1) measures the magnitude of commits made in a given month. To give a sense of scale, PyTorch requires that official maintainers successfully land 6 commits ($\log(6 + 1) \approx 1.95$) in a given technical area to be considered an expert in that area of the code base. We impute a value of 0 for both variables if no commit associated with the Contributor is observed in the relevant month.

Table 3 also provides summary statistics for two related samples of Contributor-Months. Namely, Panel B presents a strict subsample of Panel A: Non-Meta contributors to PyTorch. This sample is used in subsequent tests of Hypothesis 1 and **??**. Consistent with the idea that Meta employees are the most regular and high-intensity contributors to PyTorch, the average value of the dependent variables drops substantially relative to Panel A. Panel C presents an augmented subsample: External Company contributors to PyTorch *or* TensorFlow, a total of 964 unique contributors. This sample is used for our most stringent test of **??**. Notably, there is an extremely small number of contributors that contribute to both PyTorch and TensorFlow (29, or 2.92%), and most of them skew heavily towards contributing to either PyTorch or TensorFlow or the other. We exclude these dual contributors from Panel C and from our analysis, although their inclusion does not change the results of our analysis.

3.3 Empirical Design

To test our hypotheses, we leverage difference-in-difference style regression specifications, taking advantage of the rich data available in our setting to explore various comparison groups. All of our specifications take the following form for contributor i in month t:

$$Y_{it} = \beta_{\text{Post}} \cdot \text{Treat}_i \times \text{Post}_t + \beta_{\text{Transition}} \cdot \text{Treat}_i \times \text{Transition}_t + \alpha \cdot X_{it} + \gamma_i + \delta_t + \epsilon_{it}$$

where Y_{it} represents the dependent variable, X_{it} represents (potentially time-varying) controls that vary by specific analysis, and γ_i and δ_t are contributor and month fixed effects. Treat_i is a contributor level variable that varies by specification: for our within PyTorch specifications, it usually takes the form Chip Manufacturer_i, but for our between-PyTorch-and-TensorFlow analyses, it takes the form Chip $\text{Manufacturer}_i \times \text{PyTorch}_i$. All analyses cluster standard errors at the Affiliation and Month levels.

While the empirical specification is standard, the choice of appropriate control group for our analysis has more nuance. We only observe a single technology governance change - all contributors to PyTorch experience the same governance change at the same time. Therefore, there is no entirely unaffected group that we can leverage as a control in our analysis. Nevertheless, we explore three different reasonable comparison groups throughout our analyses, which allow us to assemble strong evidence in favor of our hypotheses.

1. *Unaffiated Contributors.* This group comprises contributors that contribute using University or Ambiguous email domains (such as 'gmail.com'). While unaffiliated contributors may well be employeed, their choice to not use their corporate email at least indicates a lower level of corporate involvement in their choice to contribute to open source software. Therefore, we'd expect that any relevant incentives passed on from a company would be lessened for this group.

2. *Non-Chip Manufacturers.* A second control group comprises the employees at cloud computing and application developer companies, such as Amazon, OpenAI, and Disney. Our theory predicts that these Users will be less affected by the governance change because their access rights to the technology do not change.

3. *External Company Contributors to TensorFlow.* A final control group is External Company contributors committing to TensorFlow, a rival machine learning framework that does not undergo the same governance change. In this case, we leverage triple-difference specifications to explore pre-post differences in the difference between Chip Manufacturers and Non-Chip Manufacturers across PyTorch and TensorFlow.

Notably, we exclude contributors from Meta from our analyses (except for our test of Hypothesis 2) because Meta initiated the governance change, and therefore Meta's contribution patterns likely anticipated and are inherently endogenous to the governance change. By contrast, we argue that the governance change was unanticipated by external company contributors prior to Spring 2022.

Before presenting these specifications, we will also present raw time trends in the data that corroborate the final statistical estimates. Further, as is standard for difference-in-difference analyses, we modify the above specification to explore how results change without fixed effects as well as the associated event study specification (to explore potential violations of the parallel trends assumptions in the form of pre-trends). Because of a lack of an unaffected comparison group, it is difficult to interpret our estimates as any direct form of average treatment effect. Nevertheless, the difference-in-difference specification allows us to tighly relate variation in control rights to variation in outcomes across different groups of theoretical interest, and for that reason we are confident in interpreting our results as supportive of a causal relationship between control rights and contributions to PyTorch. Finally, at the end of our results, we explore potential alternative explanations and associated robustness tests.

4 Empirical Analysis

4.1 Baseline Effect of PyTorch's Governance Change

Figure 1 visualizes contribution patterns across different contributor groups around the timeframe of the governance change. The top left panel shows total unique monthly contributors, revealing a striking pattern: a temporary increase in contributors during the Transition period, followed by a sharp decline after the public announcement. Breaking this down by affiliation reveals that external companies drove the initial increase, while Meta employees' participation dropped sharply post-announcement. Unaffiliated contributors showed no discernible trend throughout this period. The breakout by contributor affiliation explains this temporary increase: the increase in contributors is driven primarily by an influx of contributors by external companies, alongside a surprising temporary increase in Meta participation (potentially associated with transition-related changes). The drop-off at the end of the transition is driven by a sharp decrease in Meta contributors following the public announcement. By contrast, no discernable trend is visible for Other (Not Company). Figure A2 superimposes these time trends to illustrate the difference in trajectory depending on institutional affiliation.

To further explore this trend, the second row illustrates a sharp increase in 'exiting' contributors—those making their final commit before disappearing from the data—directly preceding the governance announcement. This pattern is primarily driven by maintainer-level Meta employees rolling off the project, with no comparable increase in exits among other contributor groups. The final row showing commit volume mirrors these patterns but exhibits higher variability.

To statistically validate these observations, we conduct regression analysis presented in Table 4, using Unaffiliated contributors as our comparison group. Model (1) confirms our visual interpretation: relative to Unaffiliated contributors, both External Company and Meta contributors show heightened activity during the Transition period. However, their behaviors diverge significantly in the Post period — External Companies maintain elevated participation levels (though with increased variability), while Meta's contribution rate becomes negative. Quantitatively, External Company contributors increased their likelihood to contribute by 25.7% ($= \frac{0.0307+0.0080}{0.0307+0.0757+.0080+.0359}$). Model (2) demonstrates these results remain robust when including fixed effects[25]. Models (3) and (4) show similar patterns when measuring commit volume rather than participation likelihood.

We cautiously interpret the results of Table 4 as evidence in favor of Hypothesis 1 — an increase in external firm participation. Although the Post-period result is not statistically significant for External Company, it is of the same magnitude as our transition period

[25] The use of fixed effects frequently does not change the point estimate in our analysis. While superficially puzzling, this fact can be explained by the observation that most contributors do not contribute in most months. (For example, as shown in the summary statistics, the average month sees participation by only 8.13% of contributors in our sample.)

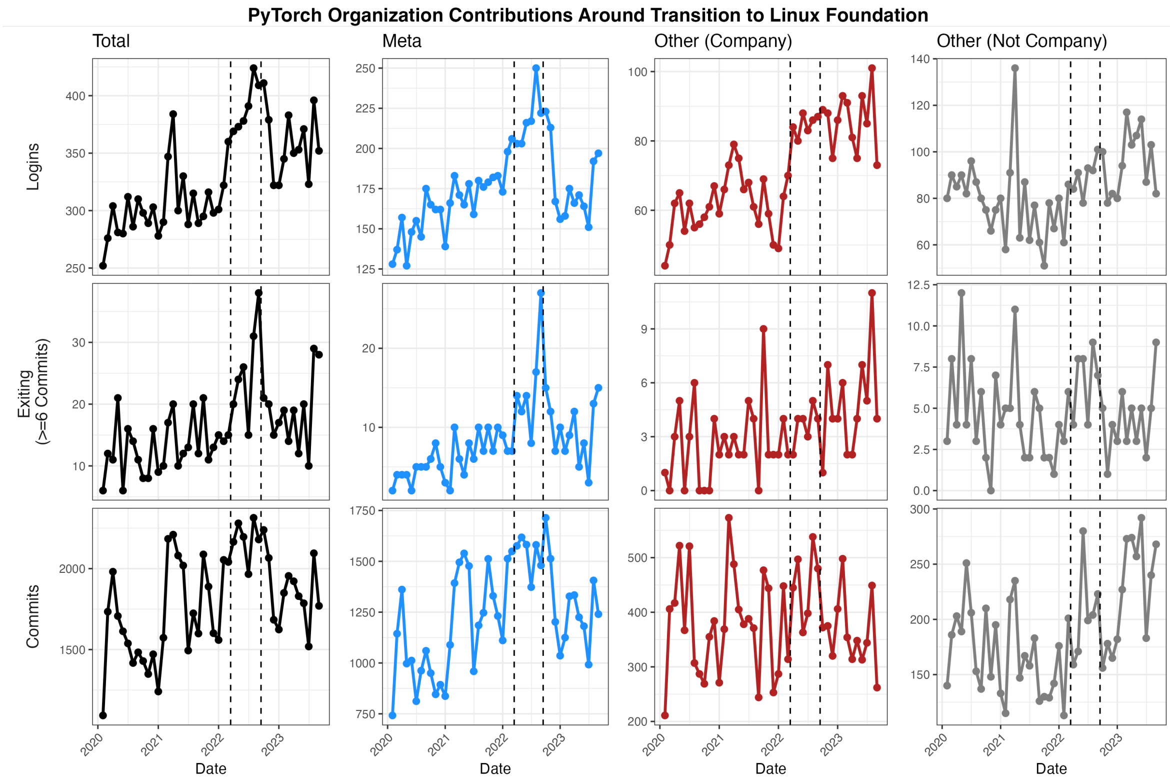

Figure 1. Time Trends in contributions to PyTorch. Each subplot shows an outcome variable (y-axis) over time (x-axis), aggregated over all repositories within the PyTorch organization on GitHub for a specific group of contributors. Each subplot row varies the specific outcome variable of interest, and each column varies the group of contributors being aggregated. In particular, the left most column presents an aggregation over all contributors to PyTorch ("Total"), and the remaining columns break out this trendline by the affiliation of the underlying contributors (Meta, Other (Company), or Other (Not Company)). Lastly, the start and the end of the Transition period (leading up to the public announcement of the Governance Change) are denoted by the vertical dotted lines. The left-most column presents the aggregate contributions to PyTorch, and the following columns split out contributions by contributor institutional affiliation type. The top row presents unique contributors contributing to PyTorch in a given month. The second row measures how many 'exits' occur in the month, where exit is defined as a contributor with ≥ 6 commits who makes their final commit in that month (and disappears from the commits data afterwards). The last row presents the aggregate number of commits made to PyTorch over the analysis period.

results and loses significance due to increased noise. Further, we emphasize that the "Not Affiliated" comparison group used in this analysis is not a perfect 'unaffected' control group — it's possible (and even likely) that the governance change has a positive effect on the contributions from this group. Therefore, the estimates for External Company are likely underestimates of the true treatment effect of the governance change. Lastly, inspection of the relevant event study plot (Figure A4) lends confidence to a causal interpretation because the are not driven by discernable pre-trends and there's a sharp and sustained increase in

Table 4. Regression analysis of the full PyTorch contribution sample (Panel A of the Table 3). The comparison group is contributors not affiliated with a known company ('Unaffiliated'). These contributors contribute with non-company emails, and come from universities or are working due to their own intrinsic motivation.

Dependent Variables:	1(Is Active)		Log(Commits+1)	
Model:	(1)	(2)	(3)	(4)
Variables				
External Company × Post	0.0307	0.0307	0.0304	0.0304
	(0.0204)	(0.0212)	(0.0317)	(0.0337)
External Company × Transition	0.0322**	0.0322**	0.0437***	0.0437**
	(0.0130)	(0.0155)	(0.0106)	(0.0212)
External Company	0.0757***		0.1291***	
	(0.0182)		(0.0351)	
Meta × Post	-0.0051	-0.0051	-0.0030	-0.0030
	(0.0031)	(0.0044)	(0.0036)	(0.0067)
Meta × Transition	0.0309***	0.0309***	0.0500***	0.0500***
	(0.0016)	(0.0041)	(0.0013)	(0.0046)
Meta	0.0850***		0.1593***	
	(0.0017)		(0.0019)	
Post	0.0080***		0.0092***	
	(0.0024)		(0.0032)	
Transition	0.0064***		0.0060***	
	(0.0018)		(0.0007)	
Constant	0.0359***		0.0338***	
	(0.0013)		(0.0015)	
Fixed-effects				
Month		Yes		Yes
Contributor		Yes		Yes
Fit statistics				
Observations	173,796	173,796	173,796	173,796
R^2	0.02759	0.27253	0.03246	0.43493
Within R^2		0.00096		0.00079

Clustered (Month & Company) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

outcome value in the point estimate at the expected time.

The results are consistent with Hypothesis 2. However, we do not interpret the results as evidence of Hypothesis 2. They are inconclusive for two reasons. First, regardless of the coefficient, we cannot give a clear causal interpretation of the result because the governance change was anticipated and is itself part of a broader technology strategy being pursued by Meta. Therefore, we cannot cleanly attribute a causal effect of the governance change to any particular time period — perhaps Meta gave up control of PyTorch because they anticipated other trends that necessitated rolling back investment. Second, while the directional results from Table 4 are consistent with the predicted causal decrease in contribution of Hypothesis 2, they are statistically insignificant given the current analysis period. However, an inspection of the associated event study (Figure A4) reveals strong upward pre-trends, which bias the results toward zero. In other words, Meta employees certainly decrease participation after the governance change, but at a similar amount to their increase in participation over the pre-period. To underscore this point, Figure A5 shows the Meta Post effect is significantly negative if we simply push the analysis start date back to May 2020 (or any date afterward)[26]. Table A2 illustrates this concretely by replicating Table 4 but only including one year of pre-period; in that case, the Meta Post result is negative and significant.

In aggregate, the preceding results highlight the implicit tradeoffs associated with changing the control rights governing a technology project. By changing from 'dominant' to 'collective' governance, PyTorch experienced an increase in contribution from External Companies. However, after a temporary increase in the Transition period, focal company (Meta) participation drops off sharply after the governance change, falling to a level comparable to or slightly lower than the pre-period. Overall, the result is that PyTorch experiences a very small increase in net sustained participation on the project, far less than would be expected based on the fanfare around the governance change. Our results highlight that, from a welfare perspective, changing governance from dominant to collective may not always increase contribution to a project because it may diminish the incentives for the focal company to continue investing. In this case, a decrease in Meta participation resulted in a much smaller increase in net contributions to PyTorch than expectations set by public messaging[27].

4.2 Increase in Company Contributions Driven by Chip Manufacturers

Why did the governance change caused an increase in External Company participation? While an interpretation of the governance change as diluting Meta's control rights and reducing the threat of hold-up for other companies is a sufficient condition for the observed increase, there could be other explanations that similarly explain the effect of the governance shift observed in Section 4.1. Accordingly, we turn to testing for heterogeneity in the effect of the governance change on the External Companies that contribute to PyTorch. By demonstrating evidence in favor of **??**, a signature prediction of our theory, we aim to demonstrate control rights as the primary mechanism behind the observed increase in external company contributors to PyTorch organization repositories.

4.2.1 Within PyTorch Analysis

Descriptively, it is clear that Chip Manufacturers started contributing significantly more to PyTorch after the Transition period began. To show this, in Figure 2, we plot the likelihood of a contributor committing to PyTorch in a given month (normalized over the pre-period values) for Chip Manufacturers, Application Developers, Cloud Providers, and Unaffiliated contributors. The plot reveals a sharp increase in Chip Manufacturer contributions around the start of the governance change. By contrast, all other categories slightly decreased their rate of participation post-governance change. Based on the similarity of the Cloud Provider and App Developer trends, we group these categories together in subsequent analysis.

To further explore this trend, we plot the descriptive contributions of the key chip manufacturers in our sample in Figure A6. The figure reveals that all but one of the chip manufacturers in the data increased their contributions closely in-sync with the timing

[26]We report the April 2020 start date results in the main paper for consistency with the rest of our results, and for greater transparency of the sensitivity of this result to the analysis start date.

[27]See https://pytorch.org/blog/one-year-pytorch/ for example.

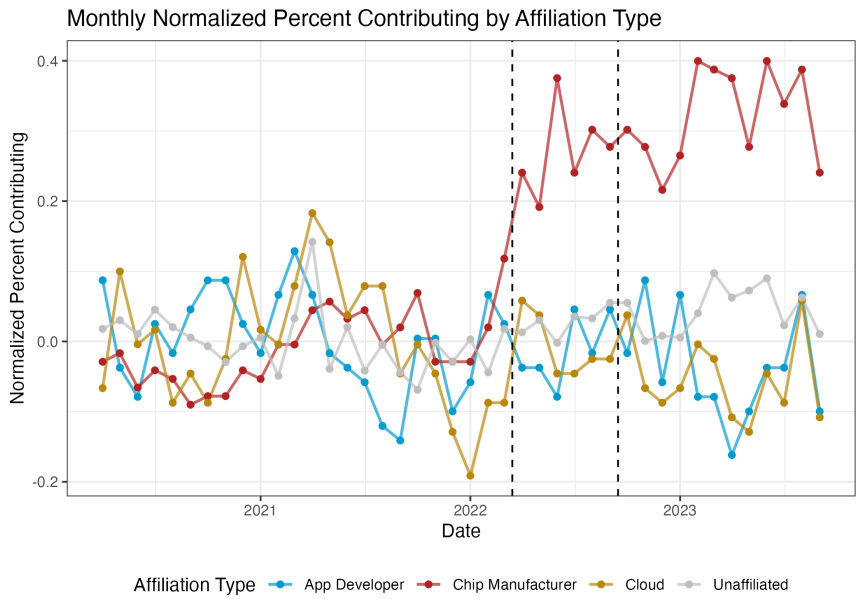

Figure 2. Trend lines of the average monthly likelihood of participation by contributors for different affiliation types, centered and scaled to normalize the pre-period values. The grey trend line depicts Unaffiated contributors. The solid red line depicts contributors employed by Chip Manufacturers. The blue and yellow lines depict App Developers and Cloud Providers (respectively). The separation of the Chip Manufacturers from other companies in the Post-period is the analytic focus of Section 4.2.

of the broader governance transition, especially Google, Intel, and Apple. The only Chip Manufacturer not increasing contributions around this time is NVIDIA, a company that had been contributing to PyTorch for years before the transition; we speculate on the reason for this curious (but tangent) fact at length in Appendix Section A.2. These quantitative findings are consistent with qualitative evidence found on the PyTorch Foundation blog, where Apple, Intel, NVIDIA, and IBM all describe the specific projects aiming to improve PyTorch efficiency on their respective hardware.[28]

While visually compelling, the above descriptive results do not tell us whether the relative increase from chip manufacturer contributors could be explained by random noise or other explanations. To statistically test this hypothesis, we explore a difference-in-difference analysis *within* PyTorch, presented in Table 5. Models (1) and (2) show a regression comparing the monthly likelihood of committing for Chip Manufacturer and Non-Chip Manufacturer Contributors versus Unaffiliated Contributors (Control Group #1), varying the use of fixed effects. The results quantify the difference shown in Figure 2, showing that Chip Manufacturer contributors dramatically increase their likelihood to contributor relative to Unaffiliated contributors in both the Transition and Post period. By contrast, Non-Chip Manufacturers decrease their relative likelihood of contribution (although this result is not statistically significant when accounting for fixed effects). Models (3) and

[28]In case the careful reader is concerned that Google is contributing to PyTorch in a capacity other than as a chip manufacturer, an examination of commits by Google reveals that almost all Google-based commits target the repository pytorch/xla – an extension focused on enabling efficient use of Google's Cloud Tensor Processing Units (TPUs) (blog).

Table 5. Regression Analysis of Non Meta Contributors to PyTorch. The unit of analysis is the Contributor-Month, and each specification presents a different choice of dependent variable, comparison group, and the inclusion of fixed effects. The results here are relevant to the test of **??**.

Dependent Variables:	1(Is Active)				Log(Commits+1)			
Comparison Group:	Unaffiliated		Non Chip Manufacturers		Unaffiliated		Non Chip Manufacturers	
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables								
Chip Manufacturer × Post	0.0939***	0.0939***	0.1151***	0.1151***	0.1090**	0.1090**	0.1438***	0.1438**
	(0.0258)	(0.0260)	(0.0273)	(0.0291)	(0.0507)	(0.0508)	(0.0517)	(0.0554)
Chip Manufacturer × Transition	0.0759***	0.0759***	0.0795***	0.0795***	0.1015***	0.1015***	0.1054***	0.1054***
	(0.0078)	(0.0137)	(0.0103)	(0.0193)	(0.0046)	(0.0159)	(0.0128)	(0.0290)
Chip Manufacturer	0.0636**		-0.0202		0.1125**		-0.0283	
	(0.0258)		(0.0360)		(0.0532)		(0.0703)	
Non Chip Manufacturer × Post	-0.0212**	-0.0212			-0.0348**	-0.0348		
	(0.0090)	(0.0132)			(0.0163)	(0.0229)		
Non Chip Manufacturer × Transition	-0.0036	-0.0036			-0.0039	-0.0039		
	(0.0038)	(0.0123)			(0.0124)	(0.0235)		
Non Chip Manufacturer	0.0837***				0.1408***			
	(0.0251)				(0.0461)			
Post	0.0083***		-0.0129		0.0089***		-0.0259	
	(0.0023)		(0.0088)		(0.0030)		(0.0160)	
Transition	0.0052**		0.0016		0.0054***		0.0014	
	(0.0024)		(0.0043)		(0.0012)		(0.0132)	
Constant	0.0358***		0.1195***		0.0338***		0.1746***	
	(0.0013)		(0.0252)		(0.0015)		(0.0462)	
Fixed-effects								
Month		Yes		Yes		Yes		Yes
Contributor		Yes		Yes		Yes		Yes
Fit statistics								
Observations	114,912	114,912	23,898	23,898	114,912	114,912	23,898	23,898
R^2	0.02872	0.19331	0.01095	0.32481	0.03679	0.36389	0.00574	0.47400
Within R^2		0.00422		0.00897		0.00430		0.00665

Clustered (Month & Company) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

(4) instead compare Chip Manufacturers with Non-Chip Manufacturers directly (Control Group #2), effectively summing the two differences observed in Model (2). Taking the point estimates from Model (3) literally, Chip Manufacturers experience a 47.1% increase ($= \frac{0.1151-0.0202}{0.1151-0.0202-0.0129+0.1195}$) in participation relative to the pre-period, compared to a -12.1% decrease ($= \frac{-0.0129}{0.1195-0.0129}$) for Non-Chip Manufacturers. Models (5)-(8) repeat the same specifications, but using Log(Commits+1) as the dependent variable. In aggregate, the results show that Chip Manufacturer employees contributions significantly increased on both the extensive and intensive margins relative to other contributor groups, providing evidence in favor of **??**. Further, inspections of the corresponding event studies (Figure A7) reveal a lack of pre-trends in the pre-period that bolsters confidence in the use of the two control groups.

4.2.2 Between TensorFlow Analysis

A central empirical concern challenging interpretation of these results is that interest in usage of AI technologies dramatically increased in December 2022 due to public release of OpenAI's ChatGPT. And while Chip Manufacturer and Application Developer companies are both affected by this demand shock, it is possible that they are differentially affected by this change in a way that confounds our analysis. To rule out this possibility, we augment our sample by further gathering the external company commits data to TensorFlow, Google's open source machine learning framework.[29] The intuition behind this expanded sample is that companies that contribute to open source machine learning frameworks (either PyTorch or TensorFlow) are subjected to the same demand forces across both technologies. Therefore, if contributions by a given company increase for PyTorch but not TensorFlow in the post-period, such an increase could not be explained by a broader shock to AI demand caused by ChatGPT.

Table 6. Regression analysis of External Company contributors to PyTorch and TensorFlow. See Table 3, Panel C for summary stats on the sample. The baseline comparison group is Non-Chip Manufacturers on TensorFlow in the Pre- period of the event. This table constitutes the primary evidence in favor of **??**.

Dependent Variables:	1(Is Active)		Log(Commits+1)	
Model:	(1)	(2)	(3)	(4)
Variables				
Chip Manufacturer × PyTorch × Post	0.0977***	0.0977***	0.1226**	0.1226*
	(0.0292)	(0.0314)	(0.0576)	(0.0618)
Chip Manufacturer × Post	0.0181	0.0181	0.0198	0.0198
	(0.0113)	(0.0132)	(0.0216)	(0.0244)
PyTorch × Post	0.0341**	0.0341*	0.0412*	0.0412
	(0.0139)	(0.0175)	(0.0226)	(0.0312)
Chip Manufacturer × PyTorch × Transition	0.0721*	0.0721*	0.1125***	0.1125**
	(0.0395)	(0.0418)	(0.0395)	(0.0507)
Chip Manufacturer × Transition	0.0103	0.0103	-0.0058	-0.0058
	(0.0298)	(0.0318)	(0.0361)	(0.0404)
PyTorch × Transition	0.0110	0.0110	0.0226	0.0226
	(0.0170)	(0.0215)	(0.0217)	(0.0353)
Chip Manufacturer × PyTorch	-0.1043**		-0.1435**	
	(0.0387)		(0.0684)	
Chip Manufacturer	0.0827***		0.1119***	
	(0.0163)		(0.0302)	
PyTorch	0.0285		0.0555	
	(0.0276)		(0.0512)	
Post	-0.0530***		-0.0753***	
	(0.0119)		(0.0195)	
Transition	-0.0154		-0.0287	
	(0.0165)		(0.0248)	
Constant	0.0949***		0.1268***	
	(0.0103)		(0.0215)	
Fixed-effects				
Month		Yes		Yes
Contributor		Yes		Yes
Fit statistics				
Observations	40,488	40,488	40,488	40,488
R^2	0.01504	0.30122	0.00987	0.41794
Within R^2		0.00888		0.00665

Clustered (Month & Company) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

We operationalize this concept by focusing on External Company contributors to *either*

[29] In the case of TensorFlow, we exclude Google contributions from our sample but retain Meta contributions; in the case of PyTorch, we exclude Meta contributions but retain Google contributions.

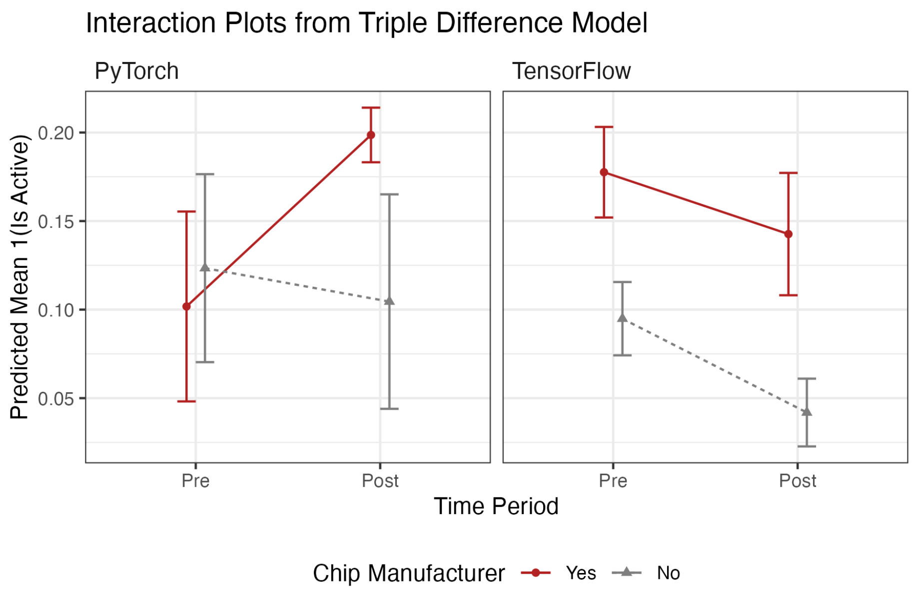

Figure 3. Interaction Effect Plot produced from Table 6, Model (1). The left visualizes the 'difference-in-difference' from PyTorch contributors, and the right visualizes the analogous effect from TensorFlow contributors. The triple-difference effect estimated in Table 6 is essentially the difference between these terms.

PyTorch or TensorFlow (see Table 3, Panel C for summary stats on this sample)[30], and augment our specification as a triple-difference analysis, presenting the results in Table 6. In this new specification, the effective 'Treatment' is Chip Manufacturer × PyTorch, and the coefficient of interest is the triple interaction term presented on the first row. We see immediately that the relative increase of Chip Manufacturers is exclusive to PyTorch: the coefficient of interest from Models (1) and (2) is positive (+0.0977) and statistically significant. Models (3) and (4) show that this effect is also found when considering the intensive margin (Log(Commits + 1)).

The magnitude of the estimate in Table 6, Model (2) (0.0977) is particularly interesting in comparison to its analog in Table 5, Model (2) (0.0939). Whereas the latter (within PyTorch) result is a difference-in-difference estimation (Chip Manufacturer vs Non-Chip Manufacturer, Pre-Post governance change), the former (between TensorFlow) result is a *triple* difference, meaning that we are observing the change in the difference-in-difference term as calculated on PyTorch and TensorFlow respectively. We infer from the fact that these coefficients are essentially equal that the equivalent difference-in-difference term for TensorFlow is therefore zero — that Chip Manufacturers did not experience a relative increase in participation on TensorFlow.

This inference is confirmed through visual inspection of Figure 3, which plots the predicted likelihood of participation for contributors to PyTorch (left) and TensorFlow (right) using Table 6, Model (1). We see that Chip Manufacturer contributors increase their likelihood of participation relative to Non-Chip Manufacturers in the Post Period within PyTorch

[30] As highlighted before, there is very minimal co-contribution to both PyTorch and TensorFlow, likely due to the significant requirements of technical skill and contextual understanding necessary for making a meaningful contribution to either of these projects. We rule out such dual contributors in our analysis.

— analogous to the result found in Table 5. In contrast, Chip Manufacturers and Non-Chip Manufacturers decrease participation at similar rates in TensorFlow (the lines are almost parallel).[31] The triple interaction term found in Table 6 essentially takes the difference of the two interaction terms plotted here: but since the TensorFlow interaction term is roughly 0, the triple-difference term from the Between TensorFlow analysis effectively equals the difference-in-difference term from the within PyTorch analysis.

As with the difference-in-difference analysis, causal interpretation of the triple-difference coefficient rests on the validity of the parallel trends assumption. And while this assumption cannot be proven empirically, we are reassured by a lack of pre-trends in the corresponding event study, provided in Figure A8.

Finally, because the above analysis assumes that company contribution to PyTorch and TensorFlow is independent, a remaining concern for this analysis is that the observed effects could be driven by competitive effects between PyTorch and TensorFlow. For example, could the increase in PyTorch contribution be driven by reallocated investment by Chip Manufacturers from TensorFlow? We argue that this is highly unlikely on the basis that we observe very few individual contributors reallocating effort to PyTorch from TensorFlow, and already filtered them out of our sample. Further, we observe Non-Chip Manufacturers similarly decline in participation on TensorFlow, but do not observe an increase in participation of Non-Chip Manufacturers on PyTorch. Finally, even if there were some competitive reallocation from companies between PyTorch and TensorFlow, we note that this reallocation would not invalidate the results from the within PyTorch analysis (Table 5), which do not require an independence assumption between contributions to PyTorch and TensorFlow.

In aggregate, we have shown that post governance change, Chip Manufacturers increase their likelihood and magnitude of contributing to PyTorch. Further, they increase their likelihood and magnitude of contributing more than Unaffiliated contributors (Control Group #1), more than Non-Chip Manufacturer contributors (Control Group #2), and more than Chip Manufacturer contributors to TensorFlow (Control Group #3). Each of these comparisons helps to rule out alternative, non-causal explanations as the reason for the observed correlation between the governance change and Chip Manfacturer participation; for example, Control Group #3 helps rule out explanations based on changes in AI demand or supply that occured around the end of 2022. The robustness of the effect across these comparisons and the lack of pre-trends gives us confidence that we are estimating a causal effect of the governance change on Chip Manufacturer's likelihood of participation and providing evidence for **??**. In particular, this heterogeneity in firm response to the governance change lends credence to our interpretation that the governance change causes this increase because of the reallocation of control rights across firms.

Consistent with this interpretation, as commented on above, Figure A6 shows that NVIDIA was the only chip manufacturer that actually *decreased* its participation post governance change, despite joining the PyTorch board. Being far and away the most regular contributor to PyTorch as well as the market leader with respect to GPUs as well as the owners of the proprietary CUDA standard (which provides a standardized interface for programming NVIDIA GPUs across different operating systems), our interpretation of

[31] Incidentally, this plot also suggests why TensorFlow is a poor group for our direct analysis from Section 4.1. Both Chip Manufacturers and Non-Chip Manufacturers were already reducing their net contributions to TensorFlow even before the governance change, in away that did not parallel their behavior on PyTorch. However, because the *difference* between Chip Manufacturers and Non-Chip Manufacturers on TensorFlow parallels that of PyTorch, TensorFlow makes a good control group for triple difference analysis.

NVIDIA's different reaction to the governance change is that NVIDIA already had significant control of the technology and had less of a threat of loss of interoperability due to the popularity of its GPUs. Speculatively, this explains why NVIDIA was one of the few chip manufacturers that did not increase its contributions significantly after the change.

4.2.3 Is the Effect Caused by Governance Board Membership?

To this point, we have assumed that the governance change uniformly affected all external firms, without accounting for specific details regarding their membership in the PyTorch governance board. However, one might be concerned that the observed increase in participation is driven by Chip Manufacturers joining the board more frequently or at different times than other firms. In other words, could it be that Chip Manufacturers are just correlated with Board Membership, and it's Board participation that drives engagement?

Descriptively, this does not appear to be the case — we plotted the contribution trends of all PyTorch governance board members in Figure A9, marking the point at which each firm joined the board. The plot clearly shows that joining the governance board does not have a uniform effect on contribution trends. For instance, Microsoft and Amazon, original board members, did not increase their contribution rate post-announcement.

To test this more rigorously, in Table A3, we expand Table 5 to include initial board membership (September 2022, which includes Meta, Google, Amazon, NVIDIA, Microsoft, and AMD) as a variable. Models (3) & (4) show that initial board membership is not associated with any change in contributions after that governance change. Models (5) & (6) interact initial board membership with whether the individual works for a chip manufacturer, and shows that it's actually the *non*-board chip manufacturers that drive the increase in post governance-change contributions. By contrast, individuals working for chip manufacturers who were on the initial board (NVIDIA, AMD) decrease their rate of participation.

These findings are consistent with the fact that board membership is optional and endogenous; firms that eventually join the board often begin contributing before joining. Based on this evidence, we interpret board membership as a signal of a firm's intent to participate rather than as a direct cause of increased contributions, justifying our analysis approach.

4.2.4 Is the Effect Caused by Substitution with Meta Contributors?

A final alternative explanation for the observed increase in firms (and in particular Chip Manufacturers) contributing to PyTorch is an alternative *causal* effect of the governance change. The literature on firm involvement in the creation of crowdsourced public goods (and OSS in particular) highlights the possibility that firm involvement can crowd out community contribution (Reisinger et al. 2014; Nagaraj and Piezunka 2024). In this setting, could the governance change have caused Chip Manufacturer entry into PyTorch simply because Meta employees stopped contributing to the parts of the code that were crucial to the integration between PyTorch and common computer chip backends? That is, perhaps it is not a shift in *control rights* but rather the *decrease in Meta participation* that drives the observed changes in Chip Manufacturer contributions, due to inherent substitutability between each firms contributions.

In order to rule out this possibility, we construct a statistical test that shows that External Companies (and in particular Chip Manufacturers) actually change contribution in a way that's positively correlated with changes in Meta employee contributions. In other words, we aim to provide evidence that Chip Manufacturer contributions are *complementary* to Meta contributions, in a way that is directly inconsistent with a crowding-out story. To do this,

1. We collect commit-level data on the specific files that were modified for each commit to PyTorch in our data set. In particular, this commit-level data features the path of every file that is modified, allowing us to granularly measure which parts of the code are being modified in the given commit.

2. Using this data, we construct "contribution groups" by grouping files in similar parts of the overall repository directory. At a high level, our goal is to identify equally-sized parts of the PyTorch codebase that are 'related', in the sense that changes to that part of the codebase are related to the same set of functionality. We accomplish this by leveraging the natural hierarchical structure of GitHub repositories, such that a contribution group is: any repository that's not the main one (`pytorch/pytorch`), any first-level subdirectory within `pytorch/pytorch` that's not in the main subdirectory (`torch/`), and any second-level subdirectory within `pytorch/pytorch/torch`. We identify 158 contribution groups using this methodology. These contribution groups capture significant differences in the type of code being created: for example, the repository `pytorch/pytorch.github.io` manages the community website, while `pytorch/pytorch/aten` manages PyTorch's abstract tensor interface.

3. We form a dataset at the contribution group level that contains the amount of contributors and commits in the six months before the governance change and the six months after (comparing October 2021-April 2022 to October 2022-April 2023). In particular, we do this for each of Meta, Chip Manufacturer, and all External Company, and compute the pre-post difference. If any employer type did not contribute to a given contribution group in either the pre- or post- period, we remove that observation from our regression (rather than marking it as zero).

We compute the correlation between the change in external company contributions and the changes in Meta contributions, presenting the results in Table 7. There, we see that Meta contributors and contributions are actually *positively* correlated with External Company and Chip Manufacturer contributions. Because we know that Meta experienced a net *decrease* in contributions over this time period, this implies that External Companies and Chip Manufacturers tended to contribute *less* in areas that Meta stopped investing in. This result is directly opposite a crowding out story: despite Meta contributions being complementary to External Company contributions, and despite a broader Meta withdrawal from contribution to PyTorch, External Company (and in particular Chip Manufacturer) participation increased. Therefore, if anything, our main results are underestimates of the causal effect of control rights on external firm participation.

4.3 Additional Robustness Checks

We conclude our analysis with an overview of four additional robustness checks, with empirical evidence presented in greater detail in the appendix.

Robustness to use of only `pytorch/pytorch`. During our discussions with the Linux Foundation, it was revealed that the Linux Foundation only took technical control of the

Table 7. Regression analysis of changes in contributions by external companies as a function of changes in contributions by Meta. The unit of analysis is the 'contribution group' (see the manuscript body for details). The variables used here are the year-over-year difference in number of commits or unique contributors to a given contribution group, by affiliation type. The results show that changes in contribution by Meta correlate with changes in contribution by external companies (including chip manufacturers), ruling out an alternative explanation involving crowding-out effects that induce entry as Meta reduces investment.

Dependent Variables:	Δ Contributors		Δ Commits	
Affiliation Type	External Company	Chip Manufacturer	External Company	Chip Manufacturer
Model:	(1)	(2)	(3)	(4)
Variables				
Δ Meta Contributors	0.1406**	0.0288		
	(0.0548)	(0.0432)		
Δ Meta Commits			0.2421***	0.1012***
			(0.0498)	(0.0368)
Fit statistics				
Observations	123	103	123	103

IID standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

main repository (`pytorch/pytorch`), rather than all 77 of the PyTorch organization's repositories on GitHub (e.g. (`pytorch/vision` or `pytorch/ignite`)). We chose to use all repositories in the PyTorch organization in our analysis because we believe that this technical detail about the governance change did not affect the overall public perception (and subsequent response to) the governance change. Nevertheless, we show in Table A4 that the use of only the primary repository (`pytorch/pytorch`) yields equivalent results.

Email Only Affiliations Robustness. As discussed above, our paper imputes corporate affiliation by augmenting email affiliations with additional information from the contributor's GitHub profiles. We show in Table A5 that our results hold if we only impute affiliation based on the email domain associated with the commit author's email.

Filter Low-Frequency Contributors Robustness. There may be a concern that our results are driven by the left tail of infrequent contributors to PyTorch or TensorFlow, who do not contribute in the majority of periods. To rule out this possibility, we repeat our analysis, but filtering to only include employees that are active for at least three months during our analysis period. Table A6 shows that our results are actually strengthened when considering this more active sub-sample of participants. That the effect seems to be driven by the intensive margin of contributor participation more than the extensive margin is consistent with the idea that strategic hold-up concerns are stronger for those already incentivized to participate, although further testing would be necessary to argue this point more fully.

5 Discussion

In this paper, we study the effect of technology control rights on an external firm's likelihood of participating in open collaboration by estimating the effect of PyTorch's governance change on external firm contributions to the project. We present two sets of findings. First, contrary to strong public expectations, the transition from governance by Meta to the Linux Foundation did not lead to a net increase in activity on the project. This is because an increase in participation by external firms was negated by a concurrent decrease in contributions by Meta. Our second set of results highlights that the increase in external firm contributors can be almost entirely attributed to increased participation by employees of Chip Manufacturers. Relative to other Application Developers and Cloud Provider company contributors or Unaffiliated contributors, these Chip Manufacturers were much more likely to begin contributing after the transition of governance to the Linux Foundation.

Why? Overall, our results are consistent with the predictions of control rights theory, which interprets the governance change as resolving a latent 'hold-up' problem that previously limited the amount of external firm participation in Meta. In particular, the first set of results highlights that governance changes do not create value out of nothing. When control rights are changed from dominant to collective, the originating focal firm loses incentives to invest. Control rights theory emphasizes that the optimal governance mode actually depends on the relative marginal benefit of effort between the focal firms and external firms, and that it is not always best to innovate under a collective governance model. The second set of results is consistent with the prediction that Chip Manufacturers ("Complementors") were subject to a greater hold-up threat than other Application Developers or Cloud Providers ("Users") because their value capture proposition depends more sensitively on technical interoperability between their chips and PyTorch. Threfore, they increased their level of participation more upon the governance transition.

Our findings contribute to the literature on value capture in open collaboration (Teece 1986; Tambe 2014; Alexy et al. 2018; Nagle 2019; Rotolo et al. 2022). Whereas the prior literature emphasizes the role of *direct* value appropriation in predicting firm participation in open collaboration (often via control of complementary assets), our paper is the first to empirically demonstrate that *future* value appropriation (through ex-post control rights) drive present-day firm participation decisions in open collaboration innovation systems. We apply a control rights framework, historically developed to answer questions about the boundary of the firm and contractual design for bilateral research agreements, to a new empirical setting: open collaboration. In doing so, we are the first (to our knowledge) to formalize and apply the concept of technology control rights as an explanation of why firms strategically open their technologies to other firms. We argue that this framework is particularly likely to apply in settings of rapid technological innovation like artificial intelligence and machine learning, where technical capabilities and, therefore, market structure are uncertain and ex-post outcomes are challenging to foresee and contract over.

Moreover, we contribute to the literature on firm participation in the governance of the digital commons (Ostrom 1990; West and O'Mahony 2008; He et al. 2020; O'Mahony and Karp 2020; Altman et al. 2022; Tang et al. 2023), which is largely qualitative or cross-sectional. Our paper contributes by studying a direct change in governance model, and provides quasi-causal estimates of its effect on external firm participation. Overall, our results go beyond *how* governance decisions are made by firms by quantitatively illustrating the economic consequences of those strategic decisions, highlighting the implicit tradeoffs, and demonstrating the importance of focal company investments when considering optimal governance mode.

Lastly, our results contribute to a literature on firm participation in OSS communities (West and Lakhani 2008; Dahlander and Magnusson 2008; Nagle 2018; Murciano-Goroff et al. 2021; Nagaraj and Piezunka 2024; Fleischmann et al. 2023; Haese and Peukert 2024; Kim et al. 2024). This literature largely focuses on firm participation in collectively governed communities or individual participation in firm-sponsored communities. However, there is a gap in our understanding of how firms contribute to OSS projects sponsored by *other* firms, a surprising gap given the likelihood of rich strategic interactions in such a setting. This paper fills this gap as one of the first studies to show what drives firm contributions in the context of another firm's (Meta's) OSS project.

5.1 Generalization

If a governance change were implemented in a similar way, should we expect a similar effect in other areas of open collaboration, such as in the context of certain platforms or other open source projects? While our analysis produces quasi-causal estimates of the effect of a governance change on external firm participation, one limitation is that a lack of a completely unaffected control group makes it hard to extrapolate these results to a more experimental setting. Nevertheless, we expect the results that we find in this study to map directionally onto many open collaboration settings.

More substantively, our theory highlights key boundary conditions on when we would expect such a shift in control rights to have an effect on external firms. Namely, control rights matter in situations where technological investment is made before noncontractable uncertainty is resolved. As argued before, noncontractability is highly likely in contexts where the regulatory environment has yet to form and businesses are still experimenting with their value capture models. This property is particularly relevant to areas of new technology development that are rapidly changing in their qualitative properties, such as in the context of Big Data and AI technology. Notably, this property is less relevant in contexts where innovation fits neatly into an existing value capture framework, such as in the context of pharmaceutical research. This observation resonates with other findings in the literature that highlight the relationship between weak patent protections and an increased prevalence of open source software (Lin and Rai 2023).

5.2 Managerial and Policy Implications

Our analysis shows the importance of control rights in encouraging entry by external firms into open collaboration. The natural, first order question that emerges is: *where do control rights come from*? The answer to this question is littered with managerial and policy implications.

On the managerial side, our results suggest that early involvement or founding of an open source project *can* be a source of control rights for a firm. In this sense, our paper suggests a new reason for the literature on firm participation in openness: by participating, firms may increase the control that they have over a technology, even if present channels for value capture are unclear. Extrapolating, our results may provide insight into why firms may invest so heavily in scientific workforces: it gives them flexibility to pivot as a novel technology takes shape in the market (Rotolo et al. 2022). Our results may also provide insight into why technical infrastructure for emergent technologies is often open: infrastructure provides a logical way to increase control rights. This potentially explains the emergence of open source machine learning hub Hugging Face and its popularity with investors despite its current lack of commercialization strategy (Greenstein et al. 2023).

For managers, the key challenge in using openness in building control rights is trading off noncontractable, future value from control rights with the fact that such control rights can deter external firms from collaborating on technology projects. Further, managers need to be able to defend the significant investments necessary to establish control rights through openness without being able to show or articulate a clear value capture channel; in this sense, this strategy is easier to implement at firms with patient executives and investors.

From a policy perspective, while retaining control rights may be an essential part of firm strategy, it also reduces third-party participation. If a firm intends to capture value through a different means (especially one that aligns value creation and value capture), then intentional concession of control rights can spur ecosystem growth. Regulators can play a key role in limiting the potential for focal firms to abuse their control rights to encourage greater collaboration in the context of open collaboration (such as in the context of the EU Digital Markets Act).

Lastly, our results highlight that regulators and open source proponents should not naively promote collective governance at all costs. The focal firm's incentives must be considered in any overall welfare calculation.

References

Philippe Aghion and Jean Tirole. The Management of Innovation. *The Quarterly Journal of Economics*, 109(4):1185–1209, November 1994. ISSN 0033-5533. doi: 10.2307/2118360. URL https://doi.org/10.2307/2118360.

Oliver Alexy, Joel West, Helge Klapper, and Markus Reitzig. Surrendering control to gain advantage: Reconciling openness and the resource-based view of the firm. *Strategic Management Journal*, 39(6):1704–1727, 2018. ISSN 1097-0266. doi: 10.1002/smj.2706. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/smj.2706. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/smj.2706.

Elizabeth J. Altman, Frank Nagle, and Michael L. Tushman. The Translucent Hand of Managed Ecosystems: Engaging Communities for Value Creation and Capture. *Academy of Management Annals*, 16(1):70–101, January 2022. ISSN 1941-6520. doi: 10.5465/annals.2020.0244. URL https://journals.aom.org/doi/full/10.5465/annals.2020.0244. Publisher: Academy of Management.

Angelique Augereau, Shane Greenstein, and Marc Rysman. Coordination versus differentiation in a standards war: 56K modems. *The RAND Journal of Economics*, 37(4):887–909, December 2006. ISSN 07416261, 17562171. doi: 10.1111/j.1756-2171.2006.tb00062.x. URL https://onlinelibrary.wiley.com/doi/10.1111/j.1756-2171.2006.tb00062.x.

Pierre Azoulay, Joshua L. Krieger, and Abhishek Nagaraj. Old Moats for New Models: Openness, Control, and Competition in Generative AI, May 2024. URL https://www.nber.org/papers/w32474.

Carliss Baldwin and Eric von Hippel. Modeling a Paradigm Shift: From Producer Innovation to User and Open Collaborative Innovation. *Organization Science*, 22(6):1399–1417, December 2011. ISSN 1047-7039. doi: 10.1287/orsc.1100.0618. URL https://pubsonline.informs.org/doi/10.1287/orsc.1100.0618. Publisher: INFORMS.

Carliss Y. Baldwin and Kim B. Clark. *Design Rules, Volume 1: The Power of Modularity*. 2000. URL https://direct.mit.edu/books/monograph/1856/Design-Rules-Volume-1The-Power-of-Modularity.

Stefano Baruffaldi and Felix Poege. Like Stars: How Firms Learn at Scientific Conferences. *Management Science*, May 2024. ISSN 0025-1909. doi: 10.1287/mnsc.2022.01373. URL https://pubsonline.informs.org/doi/full/10.1287/mnsc.2022.01373. Publisher: INFORMS.

Kevin Boudreau. Open Platform Strategies and Innovation: Granting Access vs. Devolving Control. *Management Science*, September 2010. doi: 10.1287/mnsc.1100.1215. URL https://pubsonline.informs.org/doi/abs/10.1287/mnsc.1100.1215. Publisher: INFORMS.

Henry Chesbrough and Marcel Bogers. Explicating Open Innovation: Clarifying an Emerging Paradigm for Understanding Innovation. In Henry Chesbrough, Wim Vanhaverbeke, and Joel West, editors, *New Frontiers in Open Innovation*, page 0. Oxford University Press, November 2014. ISBN 978-0-19-968246-1. doi: 10.1093/acprof:oso/9780199682461.003.0001. URL https://doi.org/10.1093/acprof:oso/9780199682461.003.0001.

Linus Dahlander and Mats Magnusson. How do Firms Make Use of Open Source Communities? *Long Range Planning*, 41(6):629–649, December 2008. ISSN 0024-6301. doi: 10.1016/j.lrp.2008.09.003. URL https://www.sciencedirect.com/science/article/pii/S0024630108000836.

Paul A. David and Joseph S. Shapiro. Community-based production of open-source software: What do we know about the developers who participate? *Information Economics and Policy*, 20(4):364–398, December 2008. ISSN 01676245. doi: 10.1016/j.infoecopol.2008.10.001. URL https://linkinghub.elsevier.com/retrieve/pii/S0167624508000553.

Fabian Fleischmann, Sven Apel, Carolin Haeussler, Patrick Figge, and Thomas Bock. Contributions of Commercial Developers in Open-Source Software – A Study of the Linux Kernel. *Academy of Management Proceedings*, 2023(1):17240, August 2023. ISSN 0065-0668. doi: 10.5465/AMPROC.2023.17240abstract. URL https://journals.aom.org/doi/abs/10.5465/AMPROC.2023.17240abstract. Publisher: Academy of Management.

Roberto Fontana and Shane Greenstein. Platform leadership and supply chains: Intel, Centrino, and the restructuring of Wi-Fi supply. *Journal of Economics & Management Strategy*, 30(2):259–286, 2021. ISSN 1530-9134. doi: 10.1111/jems.12409. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/jems.12409. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/jems.12409.

Alfonso Gambardella and Eric Von Hippel. Open Sourcing as a Profit-Maximizing Strategy for Downstream Firms. *Strategy Science*, 4(1):41–57, March 2019. ISSN 2333-2050, 2333-2077. doi: 10.1287/stsc.2018.0075. URL https://pubsonline.informs.org/doi/10.1287/stsc.2018.0075.

Cheng Gao and Rory McDonald. Shaping Nascent Industries: Innovation Strategy and Regulatory Uncertainty in Personal Genomics. *Administrative Science Quarterly*, 67(4):915–967, December 2022. ISSN 0001-8392. doi: 10.1177/00018392221112641. URL https://doi.org/10.1177/00018392221112641. Publisher: SAGE Publications Inc.

Shane Greenstein, Daniel N. Yue, Kerry Herman, and Sarah Gulick. Hugging Face: Serving AI on a Platform. *HBS Case Collection*, 2023.

Sanford J. Grossman and Oliver D. Hart. The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration. *Journal of Political Economy*, 94(4):691–719, August 1986. ISSN 0022-3808. doi: 10.1086/261404. URL https://www.journals.uchicago.edu/doi/abs/10.1086/261404.

Jeremie Haese and Christian Peukert. Commercialization of Open Source Software: Evidence from Web Browser Technologies. *Working Paper*, 2024.

Oliver Hart and John Moore. Property Rights and the Nature of the Firm. *Journal of Political Economy*, 98(6):1119–1158, December 1990. ISSN 0022-3808. doi: 10.1086/261729. URL https://www.journals.uchicago.edu/doi/abs/10.1086/261729.

Vivianna Fang He, Phanish Puranam, Yash Raj Shrestha, and Georg Von Krogh. Resolving governance disputes in communities: A study of software license decisions. *Strategic Management Journal*, 41(10):1837–1868, October 2020. ISSN 0143-2095, 1097-0266. doi: 10.1002/smj.3181. URL https://onlinelibrary.wiley.com/doi/10.1002/smj.3181.

Stephen L. Jones, Aija Leiponen, and Gurneeta Vasudeva. The evolution of cooperation in the face of conflict: Evidence from the innovation ecosystem for mobile telecom standards development. *Strategic Management Journal*, 42(4):710–740, April 2021. ISSN 0143-2095, 1097-0266. doi: 10.1002/smj.3244. URL https://onlinelibrary.wiley.com/doi/10.1002/smj.3244.

Gerald C. Kane and Sam Ransbotham. Content as Community Regulator: The Recursive Relationship Between Consumption and Contribution in Open Collaboration Communities. *Organization Science*, 27(5):1258–1274, October 2016. ISSN 1047-7039. doi: 10.1287/orsc.2016.1075. URL https://pubsonline.informs.org/doi/full/10.1287/orsc.2016.1075. Publisher: INFORMS.

Do Yoon Kim, Roberto Fontana, and Shane Greenstein. Open Devices and Slices: Evidence from Wi-Fi Equipment, January 2024. URL https://papers.ssrn.com/abstract=4718877.

Karim R Lakhani and Robert G Wolf. Why Hackers Do What They Do: Understanding Motivation and Effort in Free/Open Source Software Projects. In *Perspectives on Free and Open Source Software*. 2005.

Max Langenkamp and Daniel N. Yue. How open source machine learning software shapes AI. In *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society*, AIES '22, pages 385–395. Association for Computing Machinery, 2022. ISBN 978-1-4503-9247-1. doi: 10.1145/3514094.3534167. URL https://doi.org/10.1145/3514094.3534167.

Josh Lerner and Ulrike Malmendier. Contractibility and the Design of Research Agreements. *The American Economic Review*, 100(1):214–246, 2010. ISSN 0002-8282. URL http://www.jstor.org/stable/27804927. Publisher: American Economic Association.

Josh Lerner and Jean Tirole. Some Simple Economics of Open Source. *The Journal of Industrial Economics*, 50(2):197–234, 2002. ISSN 1467-6451. doi: 10.1111/1467-6451.00174. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/1467-6451.00174. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1467-6451.00174.

Sheen S. Levine and Michael J. Prietula. Open Collaboration for Innovation: Principles and Performance. *Organization Science*, 25(5):1414–1433, October 2014. ISSN 1047-7039. doi: 10.1287/orsc.2013.0872. URL https://pubsonline.informs.org/doi/full/10.1287/orsc.2013.0872. Publisher: INFORMS.

Yu-Kai Lin and Arun Rai. The Scope of Software Patent Protection in the Digital Age: Evidence from Alice. *Information Systems Research*, July 2023. ISSN 1047-7047. doi: 10.1287/isre.2021.0137. URL https://pubsonline.informs.org/doi/abs/10.1287/isre.2021.0137. Publisher: INFORMS.

Rory McDonald and Cheng Gao. Pivoting Isn't Enough? Managing Strategic Reorientation in New Ventures. *Organization Science*, 30(6):1289–1318, July 2019. ISSN 1047-7039. doi: 10.1287/orsc.2019.1287. URL http://pubsonline.informs.org/doi/10.1287/orsc.2019.1287.

Raviv Murciano-Goroff, Ran Zhuo, and Shane Greenstein. Hidden software and veiled value creation: Illustrations from server software usage. *Research Policy*, 50(9):104333, November 2021. ISSN 0048-7333. doi: 10.1016/j.respol.2021.104333. URL https://www.sciencedirect.com/science/article/pii/S0048733321001323.

Fiona Murray, Philippe Aghion, Mathias Dewatripont, Julian Kolev, and Scott Stern. Of Mice and Academics: Examining the Effect of Openness on Innovation. *American Economic Journal: Economic Policy*, 8(1):212–252, February 2016. ISSN 1945-7731. doi: 10.1257/pol.20140062. URL https://www.aeaweb.org/articles?id=10.1257/pol.20140062.

Elon Musk. All Our Patent Are Belong To You, 2014. URL https://www.tesla.com/blog/all-our-patent-are-belong-you.

Abhishek Nagaraj and Henning Piezunka. The divergent effect of competition on platforms: Deterring recruits, motivating converts. *Strategy Science*, 9(3):277–296, 2024.

Frank Nagle. Learning by Contributing: Gaining Competitive Advantage Through Contribution to Crowdsourced Public Goods. *Organization Science*, 29(4):569–587, June 2018. ISSN 1047-7039. doi: 10.1287/orsc.2018.1202. URL http://pubsonline.informs.org/doi/full/10.1287/orsc.2018.1202. Publisher: INFORMS.

Frank Nagle. Open Source Software and Firm Productivity. *Management Science*, 65(3): 1191–1215, March 2019. ISSN 0025-1909. doi: 10.1287/mnsc.2017.2977. URL https://pubsonline.informs.org/doi/abs/10.1287/mnsc.2017.2977. Publisher: INFORMS.

Siobhan O'Mahony and Rebecca Karp. From proprietary to collective governance: How do platform participation strategies evolve? *Strategic Management Journal*, 43(3):530–562, March 2020. ISSN 0143-2095, 1097-0266. doi: 10.1002/smj.3150. URL https://onlinelibrary.wiley.com/doi/10.1002/smj.3150.

Elinor Ostrom. *Governing the Commons: The Evolution of Institutions for Collective Action*. Political Economy of Institutions and Decisions. Cambridge University Press, Cambridge, 1990. doi: 10.1017/CBO9780511807763. URL https://www.cambridge.org/core/books/governing-the-commons/7AB7AE11BADA84409C34815CC288CD79.

Joanne E. Oxley. Appropriability Hazards and Governance in Strategic Alliances: A Transaction Cost Approach. *Journal of Law, Economics, & Organization*, 13(2):387–409, 1997. ISSN 8756-6222. URL http://www.jstor.org/stable/765097. Publisher: Oxford University Press.

Michael E. Porter. *Competitive Strategy: Techniques for Analyzing Industries and Competitors*. Simon and Schuster, June 1980. ISBN 978-1-4165-9035-4. Google-Books-ID: Hn1kNE0OcGsC.

Markus Reisinger, Ludwig Ressner, Richard Schmidtke, and Tim Paul Thomes. Crowding-in of complementary contributions to public goods: Firm investment into open source software. *Journal of Economic Behavior & Organization*, 106:78–94, October 2014. ISSN 0167-2681. doi: 10.1016/j.jebo.2014.06.005. URL https://www.sciencedirect.com/science/article/pii/S0167268114001747.

Yuqing Ren, Jilin Chen, and John Riedl. The Impact and Evolution of Group Diversity in Online Open Collaboration. *Management Science*, 62(6):1668–1686, June 2016. ISSN 0025-1909. doi: 10.1287/mnsc.2015.2178. URL https://pubsonline.informs.org/doi/abs/10.1287/mnsc.2015.2178. Publisher: INFORMS.

Joost Rietveld, Melissa A. Schilling, and Cristiano Bellavitis. Platform Strategy: Managing Ecosystem Value Through Selective Promotion of Complements. *Organization Science*, 30(6):1232–1251, September 2019. ISSN 1047-7039. doi: 10.1287/orsc.2019.1290. URL http://pubsonline.informs.org/doi/10.1287/orsc.2019.1290.

Salvador Rodriguez and Sam Schechner. Facebook Parent's Plan to Win AI Race: Give Its Tech Away Free, May 2024. URL https://www.wsj.com/tech/ai/metas-plan-to-win-ai-race-give-its-tech-away-free-4bcc080a. Section: Tech.

Alicia Rodríguez and María Jesús Nieto. Does R&D offshoring lead to SME growth? Different governance modes and the mediating role of innovation. *Strategic Management Journal*, 37(8):1734–1753, 2016. ISSN 1097-0266. doi: 10.1002/smj.2413. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/smj.2413. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/smj.2413.

Daniele Rotolo, Roberto Camerani, Nicola Grassano, and Ben R. Martin. Why do firms publish? A systematic literature review and a conceptual framework. *Research Policy*, 51(10):104606, December 2022. ISSN 0048-7333. doi: 10.1016/j.respol.2022.104606. URL https://www.sciencedirect.com/science/article/pii/S0048733322001299.

Rachelle C. Sampson. The Cost of Misaligned Governance in R&D Alliances. *Journal of Law, Economics, & Organization*, 20(2):484–526, 2004. ISSN 8756-6222. URL http://www.jstor.org/stable/3555028. Publisher: Oxford University Press.

Michael Schwarz and Yuri Takhteyev. Half a century of public software institutions: Open source as a solution to hold-up problem. *Journal of Public Economic Theory*, 12(4):609–639, 2010.

Sonali K. Shah and Frank Nagle. Why Do User Communities Matter for Strategy? *Strategic Management Review*, 1(2):305–353, June 2020. ISSN 2688-2612, 2688-2639. doi: 10.1561/111.00000012. URL http://www.nowpublishers.com/article/Details/SMR-0012. Publisher: Now Publishers, Inc.

Timothy Simcoe. Standard Setting Committees: Consensus Governance for Shared Technology Platforms. *The American Economic Review*, 102(1):305–336, 2012. ISSN 0002-8282. URL http://www.jstor.org/stable/41408776. Publisher: American Economic Association.

Prasanna Tambe. Big Data Investment, Skills, and Firm Value. *Management Science*, 60 (6):1452–1469, June 2014. ISSN 0025-1909, 1526-5501. doi: 10.1287/mnsc.2014.1899. URL http://pubsonline.informs.org/doi/abs/10.1287/mnsc.2014.1899.

Shaoqin Tang, Zhiyi Wang, and Tony Tong. Knowledge Governance in Open Source Contributions: The Role of Gatekeepers. *Academy of Management Proceedings*, 2023(1): 17622, August 2023. ISSN 0065-0668. doi: 10.5465/AMPROC.2023.17622abstract. URL https://journals.aom.org/doi/abs/10.5465/AMPROC.2023.17622abstract. Publisher: Academy of Management.

David J. Teece. Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. *Research Policy*, 15(6):285–305, December 1986. ISSN 0048-7333. doi: 10.1016/0048-7333(86)90027-2. URL http://www.sciencedirect.com/science/article/pii/0048733386900272.

Llewellyn D. W. Thomas, Paavo Ritala, Kimmo Karhu, and Mikko Heiskala. Vertical and horizontal complementarities in platform ecosystems. *Innovation*, 0 (0):1–25, 2024. ISSN 1447-9338. doi: 10.1080/14479338.2024.2303593. URL https://doi.org/10.1080/14479338.2024.2303593. Publisher: Routledge _eprint: https://doi.org/10.1080/14479338.2024.2303593.

Puay Khoon Toh and Eugene Pyun. Risky business: How standardization as coordination tool in ecosystems impacts firm-level uncertainty. *Strategic Management Journal*, 45(4): 649–679, April 2024. ISSN 0143-2095, 1097-0266. doi: 10.1002/smj.3562. URL https://onlinelibrary.wiley.com/doi/10.1002/smj.3562.

Guido Van Rossum. Origin of BDFL, 2008. URL https://www.artima.com/weblogs/viewpost.jsp?thread=235725.

Jonathan Wareham, Paul B. Fox, and Josep Lluís Cano Giner. Technology Ecosystem Governance. *Organization Science*, 25(4):1195–1215, August 2014. ISSN 1047-7039. doi: 10.1287/orsc.2014.0895. URL https://pubsonline.informs.org/doi/10.1287/orsc.2014.0895. Publisher: INFORMS.

Joel West and Karim R. Lakhani. Getting Clear About the Role of Communities in Open Innovation. *Industry and Innovation*, February 2008. URL https://papers.ssrn.com/abstract=1751048.

Joel West and Siobhán O'Mahony. The Role of Participation Architecture in Growing Sponsored Open Source Communities. *Industry and Innovation*, 15 (2):145–168, April 2008. ISSN 1366-2716. doi: 10.1080/13662710801970142. URL https://doi.org/10.1080/13662710801970142. Publisher: Routledge _eprint: https://doi.org/10.1080/13662710801970142.

Heidi L. Williams. Intellectual Property Rights and Innovation: Evidence from the Human Genome. *Journal of Political Economy*, 121(1):1–27, February 2013. ISSN 0022-3808.

doi: 10.1086/669706. URL https://www.journals.uchicago.edu/doi/abs/10.1086/669706. Publisher: The University of Chicago Press.

Nataliya Langburd Wright, Frank Nagle, and Shane Greenstein. Open source software and global entrepreneurship. *Research Policy*, 52(9):104846, November 2023. ISSN 0048-7333. doi: 10.1016/j.respol.2023.104846. URL https://www.sciencedirect.com/science/article/pii/S0048733323001300.

Nataliya Langburd Wright, Frank Nagle, and Shane Greenstein. Contributing to Growth? The Role of Open Source Software for Global Startups, August 2024. URL https://papers.ssrn.com/abstract=4699182.

A Appendix

A.1 Meta's Technology Strategy

Why did Meta invest so heavily in PyTorch in the first place, and why did they decide to spin the project out as a jointly governed foundation under the Linux Foundation in 2022? From the perspective of developing empirical tests for the effect of technology control rights, Meta's rationale for these transitions is irrelevant to this study (save for testing Hypothesis 2) as long as the change was unanticipated by external firms and unaffected by external firm strategizing. Nevertheless, we join others (e.g. see Rodriguez and Schechner (2024) or Soumith's tweet) in speculating briefly on Meta's strategic rationale here.

While Meta's overall willingness to bear a large amount of the development cost for a fully open-source tool is a puzzle, Meta has enjoyed several clear strategic advantages due to its stewardship PyTorch.[32] First, Meta clearly has significant organizational expertise in PyTorch's technology, and therefore is likely to gain competitive advantage by more efficiently leveraging the tool in their own products (like in Nagle (2018)), including taking advantage of any higher-level technologies that emerge out of the PyTorch ecosystem (such as Hugging Face, Lightning AI, or Kornia). Second, Meta enjoys a significant brand benefit that likely translates to labor market advantage in hiring talented engineers at lower wages (although empirical evidence of this is scarce). Finally, and most subtly, Meta has control of the project, at both the technical and strategic levels. PyTorch has tight technical integration with Meta's infrastructure – Meta famously runs its own production infrastructure directly from the head of PyTorch's main branch, meaning that the Github PyTorch repository is technically a mirror of the "true" Facebook-internal repo, and thus only current Meta employees can "land" pull-requests (see Eric Yang's blog or PyTorch documentation)[33]. Beyond technical control, Meta management had complete governing control over the project's technical vision and direction. This type of control is best understood as "soft power", in the sense that PyTorch prides itself on a clear division between its business and technical leadership of the project. Indeed, almost all public descriptions of the governance process highlight its democratic nature and emphasize technical, data-driven arguments over consideration of particular users or use-cases (example).

If Meta's original investment in PyTorch is economically puzzling, then its decision to spin-out the tool from official Meta control is even more puzzling. That is, what changed in Meta's incentives that led to a decision to spin-out PyTorch? Whatever the rationale, while this governance shift is plausibly exogenous to external organizations, it is unequivocally endogenous to Meta's strategic business priorities, and therefore understanding the effect of the change on Meta (beyond noting that they drastically reduce their contributions to PyTorch) is beyond the scope of our empirical analysis. We speculate this change occurred because Meta's marginal benefit of investment with singular governing control has diminished over time. As a result, the expected marginal benefit of relinquishing governing control to attract external investment became relatively attractive versus retaining control.

[32] Indeed, Meta has publicly acknowledged these advantages in shareholder calls.

[33] PyTorch leaders like Eric Yang and public documentation come off as genuinely apologetic about this arrangement, but we note that this arrangement has not changed as of 2024/03 despite at least three years of recorded public recognition of this problem.

A.2 NVIDIA's Strategic Engagement with PyTorch

While our primary analysis focuses on the implications of the governance change on overall contributions—and in particular on Meta's strategic reorientation — it is worth expanding on Chip Manufacturer's contributions to PyTorch, and the anomaly of NVIDIA's early and persistent involvement in PyTorch.

At a high level, chip manufacturers contribute to PyTorch to ensure AI developers can efficiently leverage their accelerators for AI workloads. This interoperability is essential for broad adoption and competitive performance in machine learning applications. To achieve this, manufacturers must implement core tensor operations within PyTorch's backend (ATen), developing a custom execution backend that interfaces with PyTorch's dispatcher and optimize memory management for efficient tensor processing. They also need to support various PyTorch optimizations and compiliers, including graph-based transformations via FX tracing, TorchInductor, OpenXLA, and OpenAI's Triton. While much of this development is done in-house, collaboration with the PyTorch open-source community is crucial for upstream integration, long-term maintenance, and performance optimizations. Through these contributions, chip vendors ensure their hardware remains competitive and accessible to the broader AI research and development community.

NVIDIA's substantial early contributions can be understood as part of its broader strategy to bolster its proprietary CUDA platform, which is central to accelerating deep learning computations on its GPUs. By investing heavily in PyTorch from near the project's inception, NVIDIA not only showcased CUDA's performance advantages but also helped drive adoption of a framework that was optimized for its hardware ecosystem. Indeed, NVIDIA GPU's were the only GPU with first-class support in PyTorch until PyTorch 1.8 (March 2021), which introduced an official binary package for AMD ROCm—marking the first-class support (in beta form) for AMD GPUs (Burbank 2021).

This strategic engagement contrasts sharply with the approach taken by other chip manufacturers. Whereas NVIDIA's involvement was driven by the incentives associated with a proprietary, vertically integrated technology stack, competitors such as Intel and AMD have largely promoted open standards like ROCm and OpenCL. Cross-vendor compatibility diminishes the strategic leverage that comes from close integration between hardware and software. Consistent with this, these firms have did not invest in PyTorch interoperability until much later.

Figure A6 in our appendix reveals that NVIDIA's contributions declined following the transition to a jointly governed model under the Linux Foundation. One plausible explanation for this decline is that the shift in governance diluted the unilateral influence that NVIDIA could exercise over PyTorch's technical trajectory. When control rights are concentrated within a single firm, as they were prior to the governance change, a firm like NVIDIA can more directly shape the project to maximize the benefits of its proprietary technology (in this case, optimizing PyTorch's architecture to uniquely suit the capabilities of NVIDIA GPU's [and vice-versa]). However, as governance becomes distributed among a broader coalition of stakeholders, the strategic benefits of exerting such influence are reduced. The reduced ability to steer key technical decisions may have diminished NVIDIA's incentive to sustain high levels of contribution.

From a theoretical perspective, these dynamics align with the notion of strategic hold-up: NVIDIA's early, active role can be seen as an effort to preempt potential hold-up by ensuring that PyTorch's evolution remained closely tied to the strengths of the CUDA ecosystem. Once the governance structure shifted—thus mitigating the risk of hold-up by any single dominant actor—the marginal benefits of continued intensive involvement for NVIDIA were likely reduced. In contrast, Meta's behavior reflects a different calculus, as its control over the project carried direct strategic advantages for its core business. The divergence between these patterns highlights how distinct technology strategies shape a firm's incentives in open collaboration settings.

A full exploration of NVIDIA's strategic decision-making lies beyond the scope of our empirical analysis, which focuses on why Chip Manufacturers increase in PyTorch participation *in general*. This appendix discussion offers a nuanced view of how proprietary technology standards and governance structures interact. Future research could further examine these trade-offs to provide deeper insight into the differences in incentives driving different chip manufacturers' participation in open-source projects.

A.3 Supplementary Tables and Figures

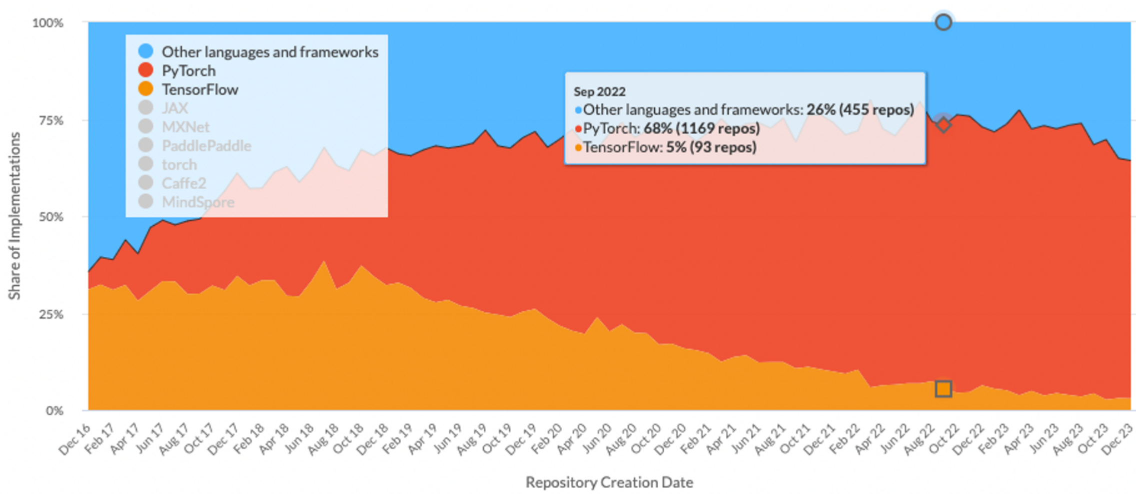

Figure A1. Increasing Dominance of PyTorch as the preferred framework for machine learning research. Figure provided by paperwithcode.com.

Table A1. Regression results demonstrating the benefit of the manually-augmented affiliation labels. The unit of analysis is unique author-email contributing to PyTorch between 2020 and 2023 for any author with an imputed corporate affiliation. While the current standard in the literature is simply to use company affiliations from email domains, an advantage of our methodology is that we add additional information from the contributor's GitHub profiles about the author's affiliation. A disadvantage is that we impute institutions at the contributor level, meaning we cannot effectively handle employment changes in our data. Usefully, our approach can be directly compared to the corresponding affiliation labels if we only used the email addresses. This allows us to quantify the informational gains from the profile information presented in this Table, by regressing the email-only imputed affiliation on the 'true' corporate affiliation imputed by using all the information that is available. Model (1) shows that Meta employees are particularly likely to use their company email address when committing rather than an ambiguous email (e.g. gmail.com). However, about 30% of chip manufacturer employees will use their Gmail or otherwise ambiguous email address affiliation. Model (2) shows that there is a mismatch of imputed affiliation on about 6.44% of unique author emails; however, Model (3) shows that it is very rare (0.61%) for this to involve an improperly imputed company, such as due to an employee move. Instead, manual inspection reveals that these mismatches come from corporate researchers using academic emails, possibly because they occurred during an internship.

Dependent Variables:	Ambiguous Email Affiliation	Different Email Affiliation	Different Company Affiliation
Model:	(1)	(2)	(3)
Variables			
Constant	0.3673***	0.0673***	0.0061*
	(0.0178)	(0.0092)	(0.0032)
Chip Manufacturer	-0.0606**	-0.0148	-0.0019
	(0.0253)	(0.0131)	(0.0045)
Meta	-0.2144***	-0.0303***	-0.0012
	(0.0194)	(0.0100)	(0.0035)
Fit statistics			
Observations	3,587	3,587	3,587
R^2	0.04297	0.00288	5.16×10^{-5}
Adjusted R^2	0.04244	0.00233	-0.00051

IID standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

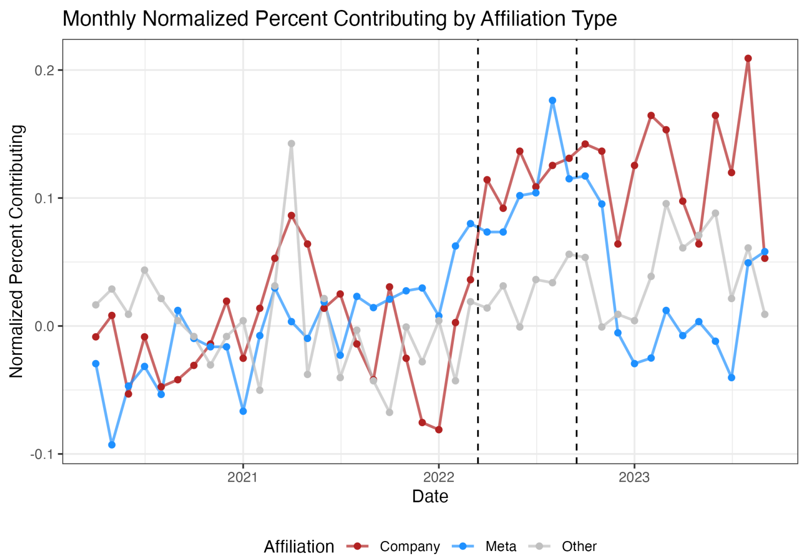
Monthly Normalized Percent Contributing by Affiliation Type
Normalized Percent Contributing
0.2
0.1
0.0
-0.1
2021
2022
2023
Date
Affiliation
Company
Meta
Other

Figure A2.

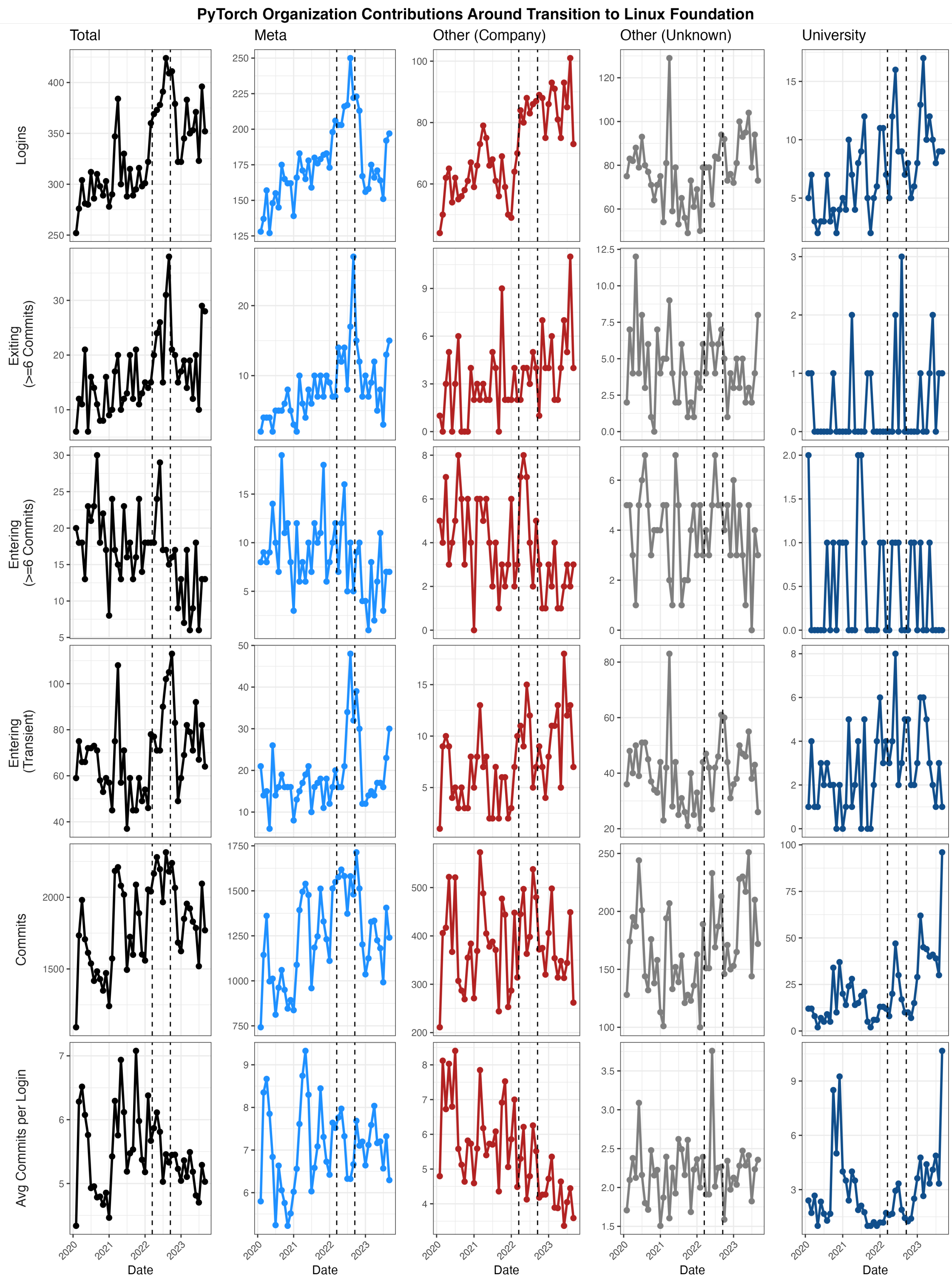

Figure A3. This figure presents an extended version of the descriptive line charts found in the main body of the paper. Here, we separate out Unknown (ambiguous) affiliations from affiliations from known universities. Further, we provide three additional outcomes: Entering (>= 6 Commits), Entering (Transient), and Avg Commits per Login. Here's its easy to see the small amount of university involvement in the project.

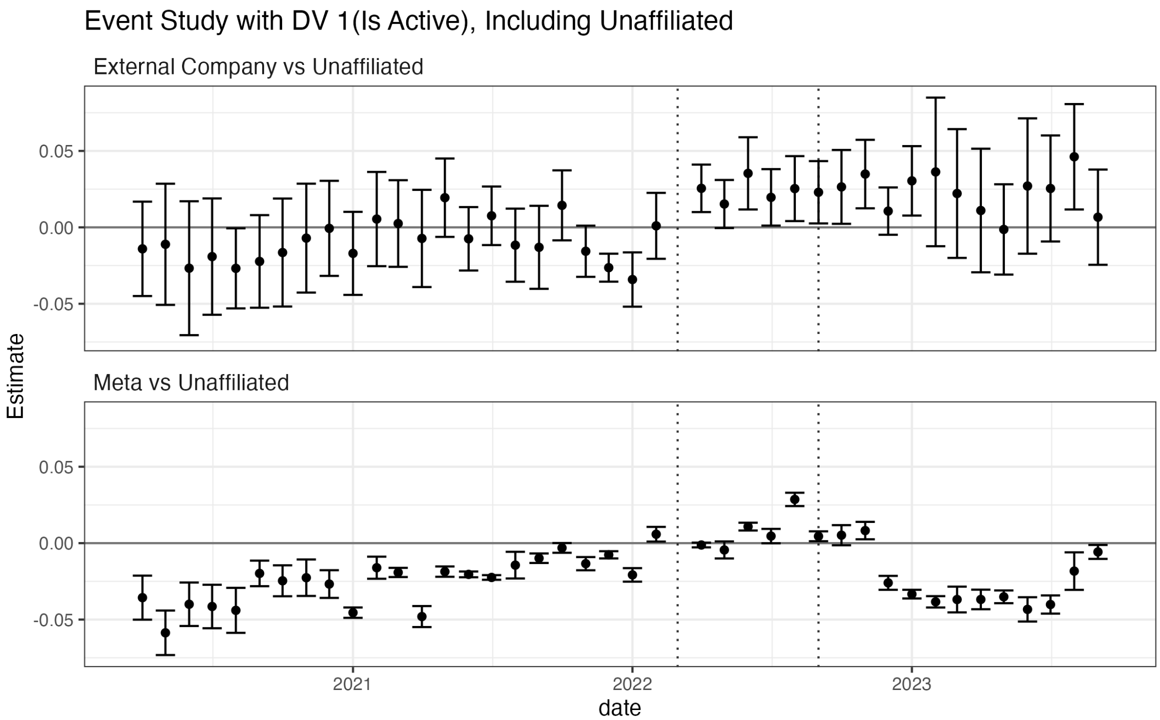

Figure A4.

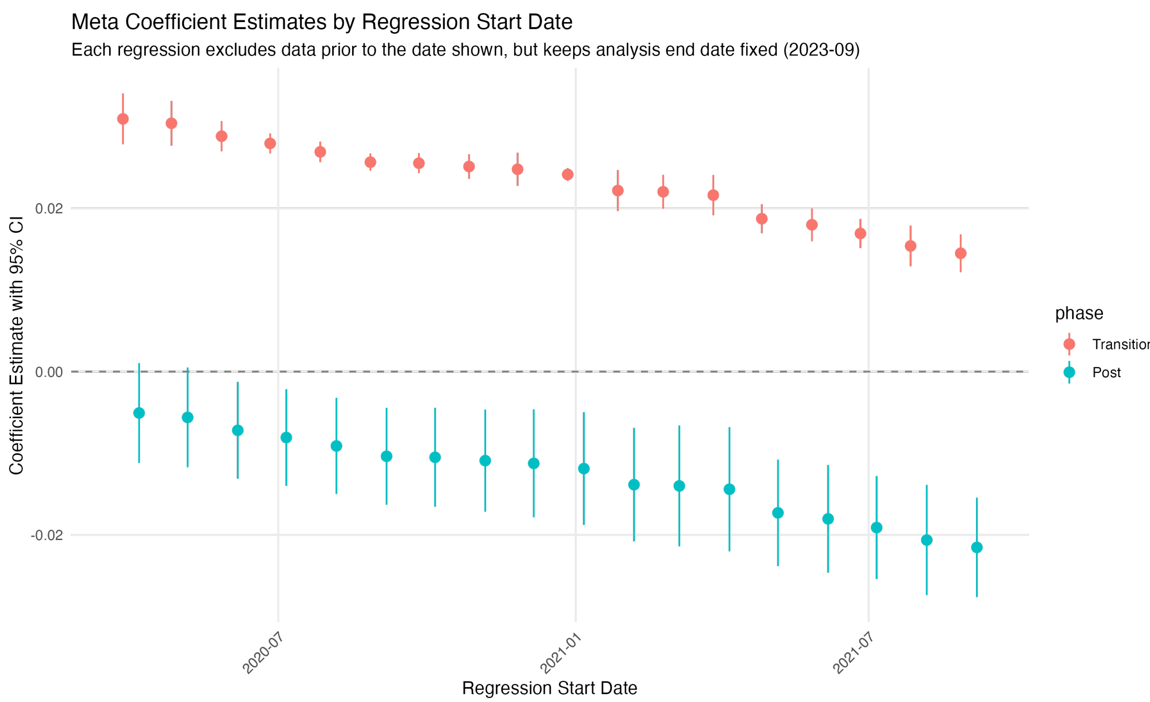

Figure A5. Coefficient Results for Meta Terms from Table 3, but with a sequentially filtered Analysis Period

Table A2. Regression analysis equivalent to Table 4, but only including one year of pre-period (April 2021 start).

Dependent Variables:	1(Is Active)		Log(Commits+1)	
Model:	(1)	(2)	(3)	(4)
Variables				
External Company × Post	0.0263	0.0263	0.0271	0.0271
	(0.0193)	(0.0207)	(0.0315)	(0.0344)
External Company × Transition	0.0278**	0.0278*	0.0403***	0.0403*
	(0.0101)	(0.0141)	(0.0025)	(0.0201)
External Company	0.0801***		0.1324***	
	(0.0193)		(0.0366)	
Meta × Post	-0.0144***	-0.0144***	-0.0213***	-0.0213***
	(0.0039)	(0.0051)	(0.0037)	(0.0070)
Meta × Transition	0.0216***	0.0216***	0.0316***	0.0316***
	(0.0013)	(0.0042)	(0.0020)	(0.0050)
Meta	0.0943***		0.1777***	
	(0.0039)		(0.0036)	
Post	0.0091**		0.0107**	
	(0.0036)		(0.0040)	
Transition	0.0075***		0.0076***	
	(0.0020)		(0.0027)	
Constant	0.0348***		0.0323***	
	(0.0040)		(0.0040)	
Fixed-effects				
Month		Yes		Yes
Contributor		Yes		Yes
Fit statistics				
Observations	124,140	124,140	124,140	124,140
R^2	0.02989	0.32200	0.03472	0.50252
Within R^2		0.00113		0.00102

Clustered (Month & aff_clean) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

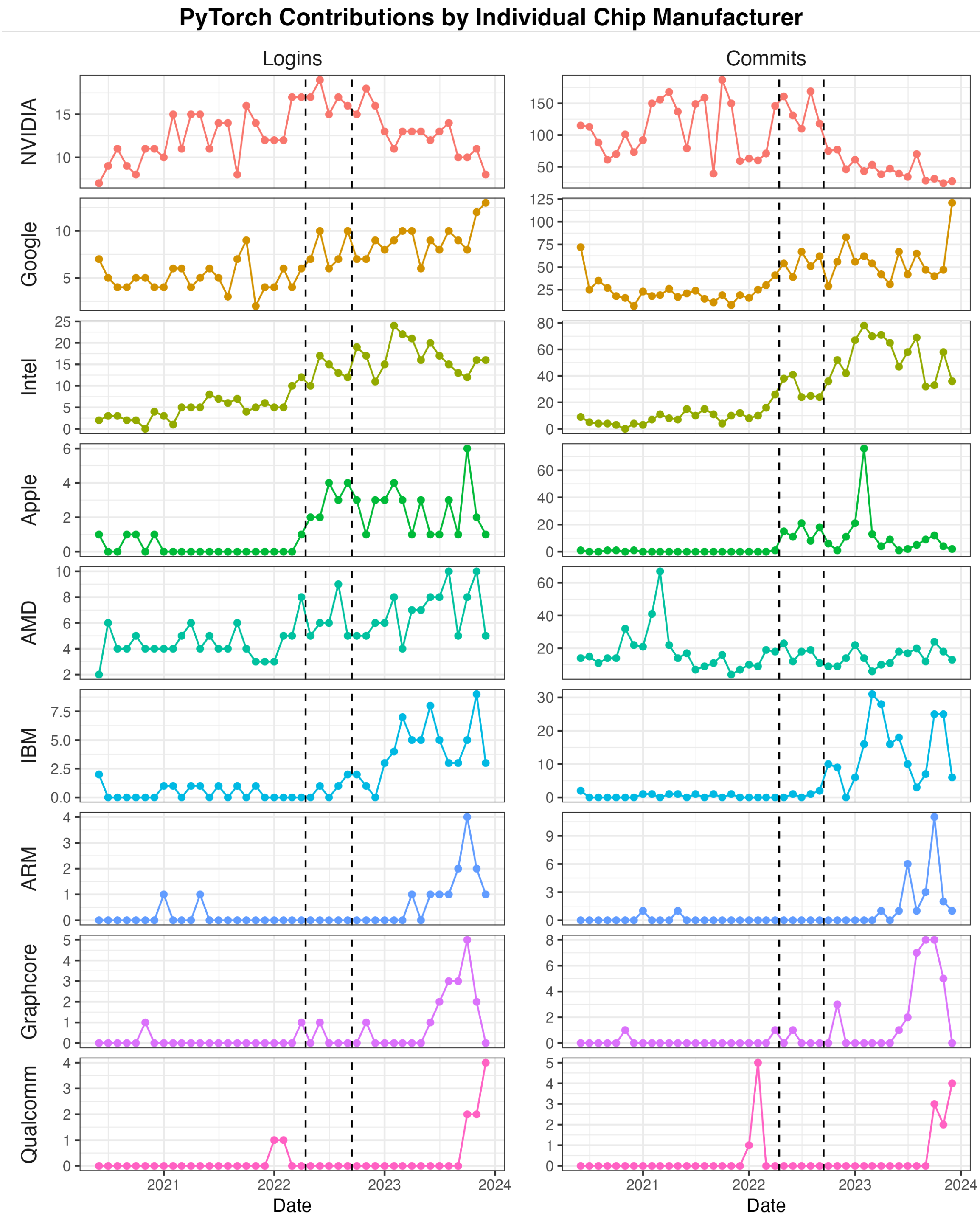

Figure A6. Individual contributions by Chip Manufacturers contributing to PyTorch organization repositories.

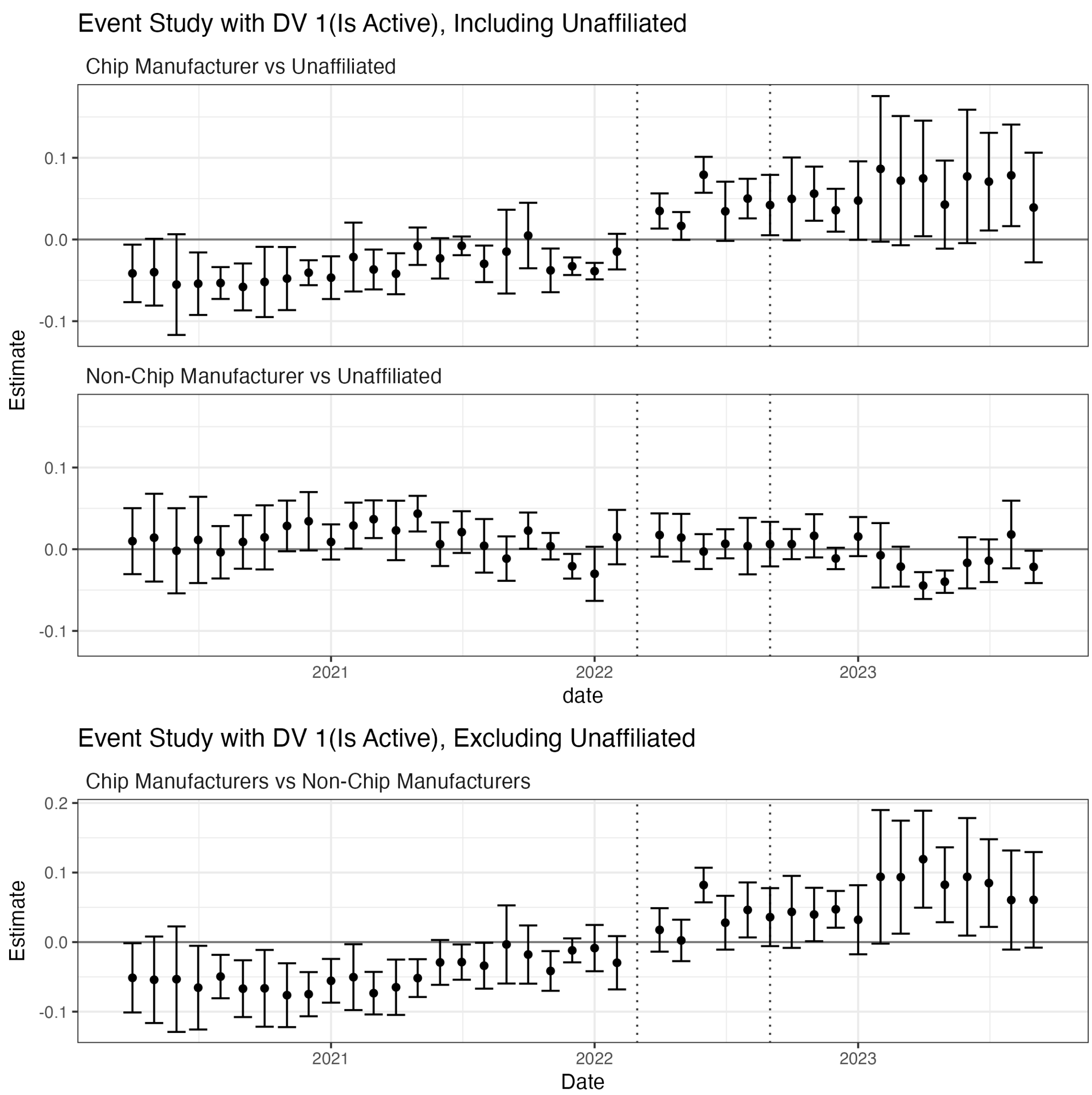

Figure A7. Event Study for Within PyTorch ComparisonsTable 5

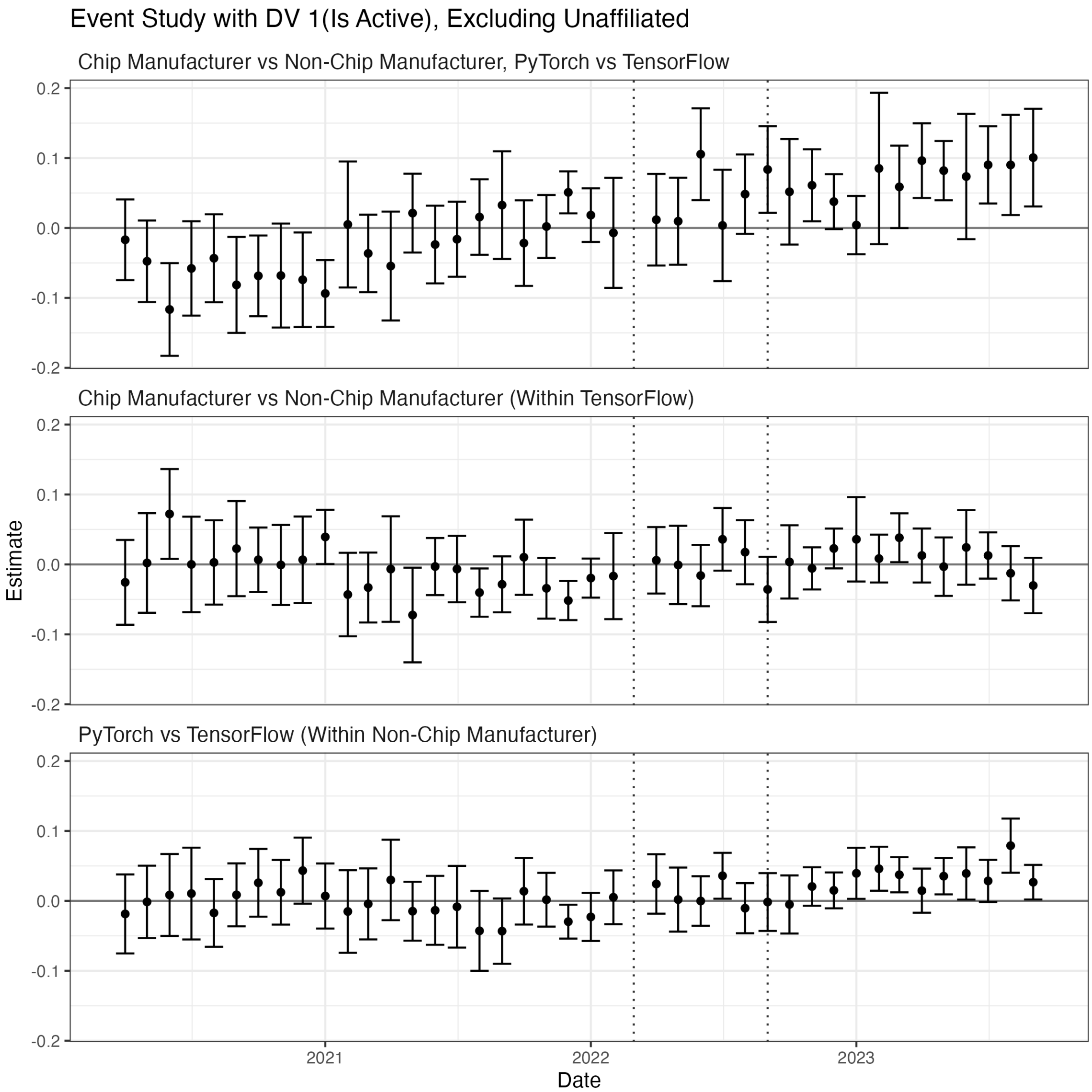

Figure A8. Event Study for Between PyTorch Comparisons Table 6

Table A3. Robustness of Results to Initial Board Membership.

Dependent Variables:	1(Is Active)	Log(Commits+1)	1(Is Active)	Log(Commits+1)	1(Is Active)	Log(Commits+1)
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Chip Manufacturer × Post	0.1134***	0.1420***			0.1640***	0.2216***
	(0.0267)	(0.0506)			(0.0196)	(0.0379)
Chip Manufacturer × Transition	0.0879***	0.1143***			0.1066***	0.1175***
	(0.0099)	(0.0130)			(0.0227)	(0.0352)
Initial Board × Post			-0.0300	-0.0689	0.0089	-0.0065
			(0.0340)	(0.0479)	(0.0160)	(0.0291)
Initial Board × Transition			0.0039	0.0064	0.0153	−0.0001
			(0.0244)	(0.0208)	(0.0092)	(0.0250)
Initial Board × Chip Manufacturer × Post					-0.1008***	-0.1552**
					(0.0272)	(0.0641)
Initial Board × Chip Manufacturer × Transition					-0.0391***	-0.0063
					(0.0140)	(0.0318)
Post	-0.0133	-0.0266	0.0547*	0.0732*	-0.0171	-0.0238
	(0.0090)	(0.0163)	(0.0302)	(0.0392)	(0.0132)	(0.0290)
Transition	-0.0032**	-0.0039	0.0369*	0.0476***	-0.0097	-0.0039
	(0.0014)	(0.0134)	(0.0203)	(0.0159)	(0.0076)	(0.0240)
Initial Board × Chip Manufacturer					0.1313**	0.2576***
					(0.0489)	(0.0901)
Chip Manufacturer	-0.0233	-0.0338			-0.0883**	-0.1608**
	(0.0359)	(0.0701)			(0.0434)	(0.0794)
Initial Board			0.0390	0.0705	-0.0220	-0.0503
			(0.0323)	(0.0632)	(0.0448)	(0.0792)
Constant	0.1212***	0.1775***	0.0920***	0.1286**	0.1306***	0.1989**
	(0.0256)	(0.0470)	(0.0284)	(0.0533)	(0.0423)	(0.0787)
Fit statistics						
Observations	23,898	23,898	23,898	23,898	23,898	23,898
R^2	0.01078	0.00560	0.00612	0.00458	0.01969	0.01874
Adjusted R^2	0.01057	0.00539	0.00592	0.00437	0.01923	0.01829

Clustered (t & Company) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A4. PyTorch Main Repo Robustness

Dependent Variables:	1(Is Active)		Log(Commits+1)	
Model:	(1)	(2)	(3)	(4)
Variables				
Chip Manufacturer × PyTorch × Post	0.0626*	0.0626*	0.0505	0.0505
	(0.0347)	(0.0369)	(0.0563)	(0.0604)
Chip Manufacturer × Post	0.0183	0.0183	0.0173	0.0173
	(0.0153)	(0.0163)	(0.0232)	(0.0251)
PyTorch × Post	0.0680***	0.0680***	0.1162***	0.1162***
	(0.0181)	(0.0219)	(0.0318)	(0.0385)
Chip Manufacturer × PyTorch × Transition	0.0531	0.0531	0.0481	0.0481
	(0.0379)	(0.0414)	(0.0384)	(0.0504)
Chip Manufacturer × Transition	0.0082	0.0082	-0.0070	-0.0070
	(0.0309)	(0.0323)	(0.0366)	(0.0397)
PyTorch × Transition	0.0187	0.0187	0.0625***	0.0625
	(0.0148)	(0.0221)	(0.0219)	(0.0379)
Chip Manufacturer × PyTorch	-0.0983**		-0.1284**	
	(0.0386)		(0.0591)	
Chip Manufacturer	0.0721***		0.0971***	
	(0.0140)		(0.0264)	
PyTorch	0.0130		0.0032	
	(0.0310)		(0.0505)	
Post	-0.0567***		-0.0818***	
	(0.0112)		(0.0189)	
Transition	-0.0145		-0.0295	
	(0.0157)		(0.0236)	
Constant	0.0901***		0.1229***	
	(0.0097)		(0.0205)	
Fixed-effects				
Month		Yes		Yes
Contributor		Yes		Yes
Fit statistics				
Observations	40,560	40,560	40,560	40,560
R^2	0.01523	0.27574	0.01185	0.36077
Within R^2		0.00919		0.00823

Clustered (Month & Company) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A5. Email Affiliations Only Robustness

Dependent Variables:	1(Is Active)		Log(Commits+1)	
Model:	(1)	(2)	(3)	(4)
Variables				
Chip Manufacturer × PyTorch × Post	0.0789**	0.0789**	0.1469***	0.1469***
	(0.0321)	(0.0331)	(0.0468)	(0.0486)
Chip Manufacturer × Post	0.0125	0.0125	-0.0052	-0.0052
	(0.0158)	(0.0167)	(0.0250)	(0.0264)
PyTorch × Post	0.0367**	0.0367**	0.0263	0.0263
	(0.0165)	(0.0174)	(0.0269)	(0.0282)
Chip Manufacturer × PyTorch × Transition	0.0418	0.0418	0.0858	0.0858
	(0.0482)	(0.0498)	(0.0563)	(0.0593)
Chip Manufacturer × Transition	-0.0005	-0.0005	-0.0323	-0.0323
	(0.0371)	(0.0390)	(0.0415)	(0.0452)
PyTorch × Transition	0.0039	0.0039	-0.0049	-0.0049
	(0.0256)	(0.0272)	(0.0388)	(0.0415)
Chip Manufacturer × PyTorch	-0.0979***		-0.1572***	
	(0.0205)		(0.0272)	
Chip Manufacturer	0.0898***	-0.1286***	0.1286***	-0.1642***
	(0.0128)	(0.0162)	(0.0213)	(0.0165)
PyTorch	0.0139		0.0338	
	(0.0144)		(0.0205)	
Post	-0.0482***		-0.0543***	
	(0.0114)		(0.0181)	
Transition	0.0013		0.0136	
	(0.0219)		(0.0275)	
Constant	0.0698***		0.0813***	
	(0.0073)		(0.0123)	
Fixed-effects				
Month		Yes		Yes
Contributor		Yes		Yes
Fit statistics				
Observations	37,632	37,632	37,632	37,632
R^2	0.01765	0.23521	0.01504	0.27101
Within R^2		0.00767		0.00801

Clustered (Month & Company) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A6. Filter Low Frequency Contributors Robustness

Dependent Variables:	1(Is Active)		Log(Commits+1)	
Model:	(1)	(2)	(3)	(4)
Variables				
Chip Manufacturer × PyTorch × Post	0.1691**	0.1691**	0.1974	0.1974
	(0.0746)	(0.0777)	(0.1469)	(0.1542)
Chip Manufacturer × Post	0.0752*	0.0752*	0.1105	0.1105
	(0.0381)	(0.0400)	(0.0680)	(0.0715)
PyTorch × Post	0.0582	0.0582	0.0925	0.0925
	(0.0491)	(0.0536)	(0.0959)	(0.1070)
Chip Manufacturer × PyTorch × Transition	0.1644**	0.1644*	0.2243**	0.2243*
	(0.0809)	(0.0850)	(0.1034)	(0.1230)
Chip Manufacturer × Transition	0.0011	0.0011	-0.0070	-0.0070
	(0.0614)	(0.0638)	(0.0906)	(0.0953)
PyTorch × Transition	-0.0114	-0.0114	0.0241	0.0241
	(0.0571)	(0.0627)	(0.0915)	(0.1109)
Chip Manufacturer × PyTorch	-0.1553**		-0.2167	
	(0.0650)		(0.1373)	
Chip Manufacturer	0.0401		0.0327	
	(0.0341)		(0.0747)	
PyTorch	0.0768		0.1424	
	(0.0521)		(0.1182)	
Post	-0.1345***		-0.2026***	
	(0.0373)		(0.0655)	
Transition	-0.0146		-0.0510	
	(0.0506)		(0.0823)	
Constant	0.2281***		0.3352***	
	(0.0311)		(0.0694)	
Fixed-effects				
Month		Yes		Yes
Contributor		Yes		Yes
Fit statistics				
Observations	17,766	17,766	17,766	17,766
R^2	0.01880	0.26798	0.01441	0.38309
Within R^2		0.01898		0.01350

Clustered (Month & Company) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

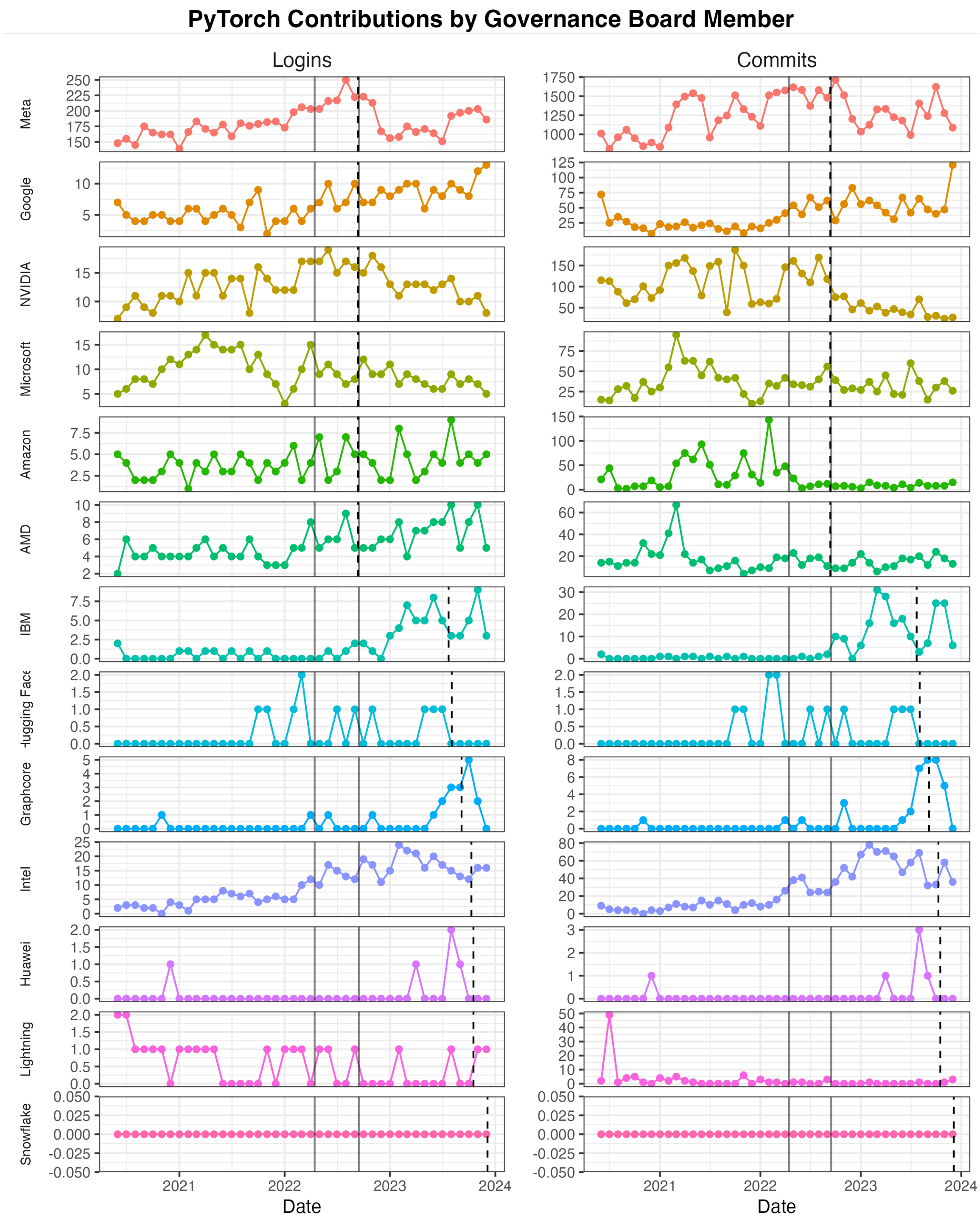

Figure A9. Individual contributions by PyTorch Governance Board members. The solid lines correspond to the dates of the Transition and Post Period. The vertical dashed line corresponds to the date that each firm joined the PyTorch Foundation Board, as announced on the public blog.