Organizational Resilience: Evidence from Open Source Software

Christopher Liao[*]

December 3, 2025

Abstract

I conduct a large-scale empirical study of open source software organizations to examine which organizational practices affect organizational resilience. Measuring resilience by the impact of a key member's exit on software development and release activity, I find that on average, organizational outcomes drop permanently by up to 0.8 standard deviations post-departure. I then develop a novel method combining flexible machine learning models with event studies that separates organizations into groups that differ markedly in measured resilience. Leveraging interpretable machine learning methods, I show that practices highlighted in the existing literature do not enhance resilience on their own, but do enhance resilience through complementarities with other practices.

1   Introduction

All organizations face challenges. Some crumble, but others are able to weather the storm. Building a resilient organization that can successfully navigate challenges is a topic of great interest.[1] Consider, for instance, the range of surveys,[2] case studies,[3] and consulting insights[4] offering guidance on building organizational resilience to disruptions like economic crises, global pandemics, or "The Great Resignation."[5] These studies identify numerous practices that are purported to enhance resilience, such as investing in talent (Maor, Park and Weddle, 2022), encouraging positive communication (Luthans et al., 2006), or establishing problem-solving routines (Suarez and Montes, 2020). A major reason for this lack of consensus is that existing research consists mainly of case studies, which focus on many different organizational contexts and disruption types.[6] This impedes structured comparisons. The relatively few recent large-scale empirical studies of the organizational determinants of resilience each measure the effect of a single aggregate shock and do not consider complementarities between different organizational practices. This underscores the need for focused, large-sample analyses that flexibly model the relationship between organizational practices and resilience.

There are two main challenges to empirically analyzing how organizational practices affect resilience. The first challenge is to clearly define and measure both organizational practices and organizational resilience. Organizations should be exposed to the same type of disruption to ensure that differences in resilience reflect differences in practices rather than differences in the disruption's nature. Resilience should be assessed using outcomes that are both comparable across yet still meaningful to all organizations. Separately, obtaining data that enables detailed, systematic measurement of resilience and practices over time and at scale is a major hurdle. Even studies leveraging rich administrative data on individuals and firms offer only limited insight into the inner workings of organizations (Jaravel, Petkova and Bell, 2018; Jäger, Heining and Lazarus, 2024).

To this end, I study organizational resilience to key member departures in open source software (OSS) organizations. OSS is a category of software that is freely available for use and modification. It is a valuable setting to study due to its economic importance: 97% of all software (including commercial types) contains OSS code (Bals 2025).[7] Understanding resilience to key member departures is also valuable because turnover is a major problem OSS organizations encounter. Separately, by studying OSS, researchers gain access to an unusually detailed picture of organizational practices and outcomes over time because activity-level data on the ongoings of many important OSS organizations is publicly available. For instance, my dataset on OSS orga-

---

[1]Robert 2009 defines resilient organizations as those able "to maintain or restore an acceptable level of functioning despite perturbations or failures."

[2]See Mallak (1998) for a survey of the health care provider industry, Danneels (2008) for a survey of U.S. Public manufacturing firms, Makkonen et al. (2014) for a survey of Finnish firms post-2008 financial crisis, and Pal, Torstensson and Mattila (2014) for a survey of Swedish textile and clothing firms during recent economic crises.

[3]See Sheffi and Rice Jr. (2005), McManus et al. (2008), Crichton, Ramsay and Kelly (2009), Dutton (2010), Lengnick-Hall, Beck and Lengnick-Hall (2011), Carmeli and Markman (2011), Suarez and Montes (2020), Tang (2025), and Atkinson (2023) for a selection of case studies and teaching cases on organizational resilience.

[4]See Maor, Park and Weddle (2022), Maor and Park (2023), and Kristensen (2025) for organizational resilience insights from McKinsey & Company.

[5]This refers to a Covid-era economic trend where employees resigned from their jobs en masse (Wikipedia, 2025)

[6]Duchek (2020) notes that "previous studies on resilience capabilities are extremely heterogeneous: They refer to different contexts, focus on specific problems, and use different research methods."

[7]Prominent examples of OSS include Linux—the operating system of choice for 96.4% of the top one million web servers (W3Cook 2015)—and the machine learning framework PyTorch—used by 63% of all organizations training machine learning models (Lawson et al., 2024).

nizations, assembled using the GitHub GraphQL API and the GitHub Archive,[8] enables me to track key members' contribution histories, measure software development activity, and characterize organizational practices.

I measure resilience by quantifying the causal impact of abrupt key member departures on software development and release activity. Focusing on individual departures allows for more precise assessments of resilience, since the disruption is clearly defined, whereas aggregate shocks often consist of multiple, concurrently evolving disruptions. For example, aggregate shocks change an organization's broader operating environment, making it difficult to separate intrinsic resilience from external conditions. Although departures are organization-specific, they are economically meaningful shocks because developer turnover is pervasive in OSS.

To identify the causal effect of the departure, I argue that the timing of abrupt departures is quasi-random and unrelated to unobservable organizational trends that may also affect software development and release activity. The literature finds that abrupt departures tend to be driven by major life events, such as a job change or starting a family, as opposed to unobservable organizational trends such as dissatisfaction with organizational leadership (Miller et al. 2019).[9]

Turning to the topic of organizational practices, I leverage my data to measure five practices that the literature finds enhance organizational resilience: *collaboration*, *knowledge level*, *discussion quality*, *investment in new talent* and *problem-solving routines*. My measure of *collaboration* evaluates how frequently members work on the same task and how evenly work is distributed. To measure *knowledge level*, I assess how experienced members are at solving many different types of problems. To measure *discussion quality*, I assess response frequency, response speed, and discussion sentiment. To measure *investment in new talent*, I assess whether organizations have infrastructure that makes integrating new members into the organization easier. And finally, to measure whether organizations have *problem-solving routines*, I assess the degree to which organizations systematize discussions, task assignments and proposed code improvements. For each practice, I assign organizations a score based on the first principal component of the constituent metrics, with higher values reflecting a greater implementation of that practice.

The second challenge is to develop an empirical framework that allows me to test hypotheses about the relationship between organizational practices and organizational resilience. Since practices may exhibit complementarities, the framework must allow a flexible functional form on the mapping from practices to resilience that supports interactions. At the same time, the framework should still be interpretable, to ensure I can discover concrete hypotheses about the relationship between organizational practices and resilience. To do this, my analysis proceeds in four steps. First, using an event study design, I estimate average resilience: the impact of key member departures on software development and release activity. To the best of my knowledge, I provide the first causal empirical evidence of these departures' effects in OSS; I show that on average, organizational software development activity is reduced by 0.8 standard deviations for at least 30 months. Second, I test whether organizations with higher practice scores are more resilient by estimating average resilience separately on organizations with high and low scores for each practice. I find that across all practices, higher scores do not suggest higher resilience.

---

[8]GitHub is the world's largest host of OSS projects. The GitHub Archive is a third-party project "to record the public GitHub timeline, archive it, and make it easily accessible for further analysis" (Github Archive 2025).

[9]Existing research from the OSS literature that studies departures broadly, without considering its abruptness or permanence, find that some departures are motivated by organizational practices, such as dissatisfaction with the organization's management or disengagement from the organization's social community (Hannon and Westlund 2008, Constantinou and Mens 2017, Miller et al. 2019). This emphasizes the importance of selecting on abrupt departures.

Since this splitting approach scales poorly to multiple practices, in my third step, I develop the *Generalized Random Forest Event Study* (separately developed in Liao 2025), which uses the tools from Athey, Tibshirani and Wager (2019) to estimate dynamic heterogeneous treatment effects in event study settings with staggered treatment (Sun and Abraham, 2021). I use the forest's predictions of organizational resilience–which are based on all pre-departure practice scores and their interactions–to construct the *Forest Split*, which separates out-of-sample organizations into equally sized high and low resilience subsets. I find that the Forest Split has very high predictive power: on average, low resilience organizations experience declines between 0.79 and 2.15 standard deviations larger than high resilience organizations across all outcomes measuring software development and release activity. Finally, in the fourth step, to aid interpretation, I coarsen the organizational practice space to reduce the dimensionality of the forest's model of predicted organizational resilience conditional on those practices. The lower-dimensional results suggest complementarities among practices: although high collaboration scores alone do not predict greater resilience, organizations with high collaboration scores are more resilient when they also have high member knowledge scores or high discussion quality scores.

As the *Forest Split* accurately predicts the resilience of out-of-sample organizations, I am confident that important features of the true statistical distribution of organizational resilience, conditional on the organizational practices in my analysis, are captured. One challenge, however, is translating these statistical relationships into causal claims. The adoption of practices is likely endogenous to unobservable factors that also affect resilience. Consequently, my findings about the statistical relationship between practices and resilience should be viewed as descriptive rather than causal; establishing which practices are causal drivers remains an important avenue for future research.

The remainder of the paper proceeds as follows. The Related Literature section situates this paper within the economics, management, and software engineering literatures. Section 2 provides background on open-source software development, describes the dataset, and defines and measures key members and abrupt departures. Section 3 describes the organizational practices studied in this paper. Section 4 leverages abrupt key member departures to estimate average organizational resilience and tests how average resilience is affected by individual practices. Section 5 uses machine learning methods to discover whether more flexible hypotheses can better explain the relationship between practices and resilience. Section 6 interprets the machine learning models from Section 5 to learn hypotheses about how practices affect resilience. Section 7 concludes.

Related Literature

A large literature in economics has studied how organizations and teams respond to the plausibly exogenous departure of different types of members, such as superstar research collaborators (Azoulay, Graff Zivin and Wang, 2010; Azoulay, Fons-Rosen and Graff Zivin, 2019), co-inventors (Jaravel, Petkova and Bell, 2018), co-workers (Jäger, Heining and Lazarus, 2024), coauthors (Oettl, 2012; Khanna, 2021), and graduate advisors (Waldinger, 2010). However, even studies that leverage large-scale administrative data to study the effects of plausibly exogenous departures (Jaravel, Petkova and Bell, 2018; Jäger, Heining and Lazarus, 2024) offer limited insight into the internal workings of organizations and, consequently, into how organizational practices shape resilience. My paper's contribution to this literature is twofold: I show that the departure of key software developers has large, negative effects on OSS organizational software development and I leverage rich data on OSS organizations to discover hypotheses about the relationship between organizational practices and organizational resilience to departures.

The literature on management and firm performance shows that adopting management practices contributes to improved outcomes (Bloom et al., 2012; Bloom, Rafaella and Van Reenen, 2017; Bruhn, Karlan and Schoar, 2018). A study of Italian firms during COVID-19 finds that the adoption of management practices also contributes to resilience; organizations with higher management practice scores had higher sales during the pandemic (Lamorgese et al., 2024). This paper's contribution is to examine how multiple organizational practices—and their combinations—suggest enhanced organizational resilience, as opposed to the effect of just a single management practice score. A related theoretical (Milgrom and Roberts, 1990, 1995) and empirical literature (Ichniowski, Shaw and Prennushi, 1997) emphasizes that organizational practices benefit from complementarities. Organizations benefit from adopting combinations of practices, not individual practices. This paper's contribution is to show that these complementarities also shape resilience, including outside of "normal times." High collaboration scores do not suggest increased organizational resilience on their own, but combined with higher discussion quality or knowledge level scores, they do. This paper also contributes to the longstanding debate over the existence of universal "best practices." I show that high and low resilience organizations can be differentiated by their pre-departure organizational practices, lending support to the universal "best practices" perspective (Taylor, 1911).[10]

My third contribution is to the literature on OSS within economics that began with Lerner and Tirole (2001, 2002). Recent empirical work has focused on the relationship between the public and private sector and OSS organizations.[11] This paper focuses on just OSS organizations, leveraging the visibility of their internal dynamics to study how an important problem OSS organizations face can be mitigated: disruption caused by key member departures. I build on prior empirical work that uses detailed activity-level data on OSS organizations to characterize organizational outcomes by also using these data to characterize organizational practices.

A large body of conceptual and empirical research on the organizational practices[12] that contribute to organizational resilience exists in the management literature, although most empirical work relies on case studies or surveys.[13] Three central challenges to empirical research on organizational resilience are in defining it, conceptualizing measurement, and testing how organizational practices affect resilience. I define organizational resilience as the capacity of organizations to respond to disruptions or shocks.[14] I measure organizational resilience by analyzing organizational outcomes observed following the disruptive event. The five categories of organizational practices in my analysis are *collaboration*, *knowledge level*, *discussion quality*, *investment in new talent* and *problem-solving routines*. These categories map to common themes emphasized by the organizational resilience and have all been found to bolster resilience.[15]

---

[10]This contrasts with the "contingency view," which says there is no single best way to organize a corporation (Woodward, 1965).

[11]Nagle (2017) and Conti, Peukert and Roche (2025) study firm gains from OSS contributions. Conti et al. (2023) and Gortmaker (2025) study the role of financial sponsorship and industrial policy in incentivizing OSS innovation.

[12]Also described as "factors" and "processes" in the literature.

[13]See review papers on organizational resilience by Annarelli and Nonino 2016 and Hillmann and Guenther 2021, and Table 2 of Annarelli and Nonino 2016 for a breakdown of methods used in studies of organizational resilience.

[14]This definition is one of several established definitions in the literature. Other work from the literature that has also used this definition include Home III (1997), Kantur and İşeri Say (2012), Boin and van Eeten (2013), and Sahebjamnia, Torabi and Mansouri (2015). Other definitions used by the literature include an organization's ability to respond to adversity, crisis, stress, turbulence, a competitive environment or unexpected events. See Table 2 of Hillmann and Guenther (2021) for papers associated with each definition.

[15]Three of the four concepts that Duchek (2020) describes as "general attributes that may facilitate an organization's resilience" cover my categories. What Duchek (2020) terms *redundancy* covers collaboration and

Existing empirical research in management examines individual organizational practices in isolation or the combined effects of multiple practices without considering complementarities between practices. For example, Ortiz-de Mandojana and Bansal (2016) and DesJardine, Bansal and Yang (2019), which examine the relationship between a company's social and environmental practices and post-financial crisis performance, but not interactions between practices. When interactions are examined, they typically involve how financial resources enable other organizational practices, such as positive labor relationships (Gittell et al., 2006), which falls outside my focus.[16] Existing research on complementarities in organizational resilience is largely conceptual and even then, does not examine in detail how practices interact to contribute to resilience. Kendra and Wachtendorf (2003) describe how knowledge level, routines, and investing and incorporating new talent and resources all contribute to redundancy but do not examine their interdependence within the system. Other work, such as McKinsey's organizational resilience insights, emphasizes building multiple capabilities simultaneously—like self-sufficient teams (that collaborate) or investing in positive relationships (termed culture) and new talent—but does not examine how these practices complement one another (Maor, Park and Weddle, 2022).

What makes OSS organizations resilient post-departure is also an active research topic in the information systems and software engineering literature. Rashid, Clarke and O'Connor (2019) provides a detailed review. There has been extensive conceptual, survey, and case study work on the prevalence of departures and how OSS organizations can mitigate either the probability of member departures or their impact.[17] Rashid, Clarke and O'Connor (2019) highlights disaster-mitigation practices that can be classified into my five categories of organizational practices.[18] One outcome prior empirical work uses to quantify the impact of departures is "potential damage to the codebase" (Izquierdo-Cortazar et al. 2009, Rigby et al. 2016, Nassif and Robillard 2017). Rigby et al. 2016, which is most closely related to my paper, finds that the presence of "successor" who worked on tasks similar to the departed member, can mitigate the repercussions of departures. My analysis differs in several ways. First, I focus on a variety of organizational practices, as opposed to just one. Next, while Rigby et al. (2016) analyzes all departures from just 5 organizations, my analysis includes 681 organizations. Finally, while Rigby et al. (2016) considers hypothetical impacts based on potential codebase damage and possible mitigation through a successor, I quantify, using realized outcomes of organizational output, the actual impact of organizational practices.

2 Background and Data

My analysis focuses on organizational resilience in OSS organizations for four reasons. First, rich data on activity and outcomes enables me to identify which organizational practices—or combina-

---

knowledge level, *positive relationships* covers discussion quality and investment in new talent and *specific organizational strategies* covers organizational routines. The fourth type is adequate resources, which is less applicable to OSS organizations as most organizations are made up of volunteers who collaborate remotely.

[16]Since many OSS projects are distributed for free, financial resource management is less relevant than in traditional corporate organizations.

[17]See work in this area by von Krogh, Spaeth and Lakhani (2003), Robles, Gonzalez-Barahona and Michlmayr (2005), Hannon and Westlund (2008), Xu and Jones (2010), Yu, Benlian and Hess (2012), Rashid, Clarke and O'Connor (2017), and Miller et al. (2019)

[18]See section 6.2 of Rashid, Clarke and O'Connor (2019). The collaboration practice is "pair programming and shared code ownership." The knowledge level practices are "successor[s]," and "uniform knowledge distribution." The discussion quality practice is "improving code review feedback time for non-cores." The investment in new talent practice is "removal of knowledge barriers." The problem-solving routine practice is "gamification."

tions of practices—suggest enhanced resilience. Second, OSS organizations are comparable: they follow similar processes and pursue the shared goal of producing new and updated software. Third, key member departures offer a natural setting for studying resilience because developer turnover is pervasive and a major challenge OSS organizations encounter. Finally, OSS—and the digital economy it underpins—are economically significant: firms would need to spend an estimated 3.5 times more on software without OSS (Hoffmann, Nagle and Zhou, 2024), and U.S. software spending already exceeds 1% of GDP (Bonakdarpour et al., 2021).

Section 2.1 describes the OSS landscape present in my sample and my data sources. Section 2.2 describes the OSS software development process and organizational structure of OSS organizations. Section 2.3 describes my procedure for identifying an organization's key members. Section 2.4 describes how I define and measure abrupt key member departures, and discusses why they can be considered plausibly exogenous. Section 2.5 describes the final analysis sample.

2.1 Sample Universe and Data Sources

I analyze an important subset of the OSS ecosystem—widely used Python libraries—from 2015–2024.[19] Python is currently the world's most popular programming language (Jansen 2025), and its libraries—collections of reusable toolkits providing prewritten code for specific tasks—are a central reason for its popularity. Most Python library development occurs on GitHub, a cloud-based platform for software development and version control. The unit of observation in my analysis is the organization, defined as the GitHub project responsible for developing one or more Python libraries.

I assemble a dataset of activity-level data on these organizations from the GitHub Archive (Github Archive, 2025), supplemented by the GitHub GraphQL API (Github, 2025b) and the Python Software Foundation (The Python Software Foundation, 2025). My dataset includes activity-level information on software development, including members' deanonymized identities and the date and content of their contributions, which I aggregate to the six-month level.

## 2.2 Background on Open Source Software Development

On GitHub, members initiate discussions (for example, about bugs or feature requests) by opening an *issue* thread and participate via *issue comments*. To propose changes to the codebase, members open a *pull request* (PR) thread. Pull request *reviewers* discuss broader aspects of the proposal through *pull request comments* and address specific code changes through *review comments*, before making determinations through *reviews*. If the *reviewers* approve the proposed changes, the pull request is *merged* into the codebase. Merged changes become available to users only when the organization publishes a new *software release*, which bundles one or more PRs and provides an updated version of the library. Figure 1a depicts the process that an OSS organization undergoes to produce a new software release.

I use the umbrella term *discussion thread* to include issue and pull request threads. I classify opening and merging PRs as software development activity and releasing new software as software release activity. Software releases are less frequent than opening or merging pull requests because they usually bundle multiple changes. While pull request activity best captures an organization's general level of activity, users tend to care more about releases, which deliver features and bug fixes that directly affect them.

---

[19]I define a Python library as widely used if it had at least 10,000 downloads from pip in any month between 2015-2024, as measured by The Python Software Foundation (2025)

Figure 1: OSS Development Workflow and Social Network

Panel (a): Release Production Workflow
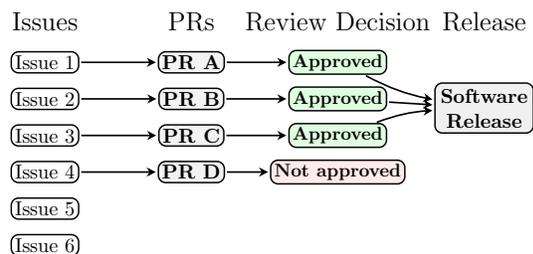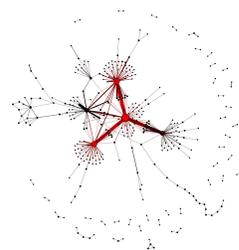

Panel (b): Social Network for `boto/boto3`

Figure notes: Panel 1a depicts the development process for an OSS to produce a new software release. Panel 1b depicts the social network for `boto/boto3`, the Amazon Web Services Software Development Kit for Python, during the first half of 2020. Each node represents a member and is scaled by its degree, defined as the number of members they interact with. Each edge between nodes represents interactions between members; edge thickness is positively related to the quantity of interactions between members. The three nodes representing the three *single-period key members* and their corresponding edges are colored in red.

I consider anyone who participates in these activities to be a member of the OSS organization. Prior research finds that a small minority of key members account for most activity (Mockus, Fielding and Herbsleb, 2002; Crowston and Howison, 2006). Key members are persistent, long-term contributors and interact broadly with others to discuss software-related topics and make core codebase decisions (Mockus, Fielding and Herbsleb, 2002; Bao et al., 2021). In contrast, non-key members usually engage only when they encounter issues relevant to their own use and often stop participating once their problems are resolved (von Hippel and von Krogh, 2003).

2.3 Identifying Key Members

Identifying an organization's key members requires a measure of each member's prominence. I quantify prominence using the organization's social network, modeled as an undirected, weighted graph whose nodes represent members and whose edge weights capture the number of interactions between them.[20] Each graph is built from six months of activity, representing either the first or second half of the calendar year. This network-based measure has two advantages: it aggregates interactions across all discussions and it draws on established graph-theoretic notions of centrality to identify the most prominent members. My definition of an OSS organization's social network follows Crowston and Howison (2006).

A member's prominence is defined by their degree centrality: the number of unique members they interacted with during the period. Members ranking among the top three in degree centrality are *single-period key members*; all ties are included.[21] Members who are single-period key members for three consecutive periods are classified as *key members* beginning in the third period. The criteria for identifying key members reflect two established features known from the literature on OSS: key members interact with more people than others and tend to have persistent contribution patterns. Figure 1b depicts the social network and single-period key members for *boto/boto3*, the Amazon Web Services Software Development Kit for Python, during the first half of 2020.

---

[20]An interaction occurs when one member's comment in a discussion thread follows another member's post or comment.

[21]More formally, members whose degree centrality equals or exceeds that of the third-ranked member are included.

2.4 Abrupt Departures

An abrupt key member departure occurs when a key member ceases all involvement with an OSS organization. I classify a key member in one period as having abruptly departed in the next if they make no contributions in any subsequent period and were a key member in the current period. This measure builds on Miller et al. (2019), who study abrupt departures using survey and contribution data.[22]

2.4.1 Motivations for Abrupt Departures

My analysis focuses on abrupt departures because survey evidence shows they typically arise from "some kind of transition (e.g., switching jobs or leaving academia)" or because members were "having children or getting married" (Miller et al., 2019). These transitions often result, by construction, in disengagement from OSS. For example, contributors whose OSS activity was tied to graduate school or to a particular job often cease contributing after graduation or when changing employers.

Abrupt departures should be distinguished from other forms of disengagement because the underlying motivations differ. For example, Miller et al. (2019) find that

> occupational reasons such as major life changes (e.g., getting a new job or leaving school) were the most cited (with 106 citations), significantly more than lacking peer support or losing interest that are more commonly discussed in the literature (Miller et al., 2019).

Outside of this study, most work examines disengagement broadly, focusing on reduced activity or gradual withdrawal rather than the speed of the exit. Commonly cited reasons in that literature include lacking peer support, role changes, and declining interest (Iaffaldano et al., 2019).

My definition of abrupt departures excludes departures with these latter motivations. Social factors are often cited as barriers to joining or becoming integral to a project (Bosu and Carver 2014; Steinmacher, Treude and Gerosa 2019). However, my analysis focuses on the departure of key members who have been integral to a project for at least 18 months. Role changes may lead members to reduce or cease code contributions, but since my definition requires complete disengagement, role shifts involving continued participation—such as supervisory or discussion-based roles—would not qualify as departures. Finally, loss of motivation caused by declining interest or rising frustration is more likely to manifest as a gradual decline in activity among long-term contributors, not an abrupt stop.

2.4.2 Plausible Exogeneity of Abrupt Departures

I argue that the timing of abrupt departures is plausibly exogenous and unrelated to unobservable organizational trends that affect organizational outcomes. This assumption would fail if departures were systematically driven by organizational factors such as dissatisfaction with leadership, lack of peer support, or declining project relevance.

However, as discussed in Section 2.4.1, the primary motivations for abrupt departures stem from individual life transitions—graduation, job changes, and family events—that are external

---

[22]My definition differs from Miller et al. (2019) by identifying key members using all forms of activity—not only code contributions—and by defining departure as a complete cessation of activity rather than a substantial reduction.

to the OSS organization. These transitions determine the timing of departure independently of unobserved organizational trends. For instance, when an OSS member has a child, the resulting reduction in available free time, which manifests as OSS disengagement, is unrelated to changes within the OSS organization. Similarly, when contributors graduate or change jobs, disengagement follows from that transition. Job changes, in particular, are usually driven by compensation, career progression, or relocation—factors orthogonal to the OSS organization's trajectory, Even when contributors use the software at work, job transitions need not reflect trends in the OSS organization. Some employers restrict OSS participation during work hours, and contributors who switch firms may stop contributing simply because the new job no longer requires or permits it. When an employer itself manages the OSS organization, the software typically represents a minor or non-revenue-generating part of the firm's operations, making it unlikely that unobserved OSS-level shocks drive or determine the timing of job exits. Hence, it is also unlikely that a decline in the OSS's popularity would lead employees to depart or firms to dismiss them.

Taken together, these patterns suggest that abrupt departures primarily reflect idiosyncratic individual shocks rather than responses to latent organizational trends. As such, I treat their timing as plausibly exogenous with respect to the unobserved organizational trends affecting OSS organizational outcomes.

2.5 Defining Treatment and the Final Sample

My sample spans 10 years (2015–2024), divided into semiannual time periods. Treated organizations experience exactly one key member departure during this window; control organizations never do. Organizations with more than one key member departure are excluded to avoid complications from multiple treatments.

I define the treatment date E_i for an organization i as the last period t in which the departing key member is present. For control organizations that never experience a key member departure, $E_i = \infty$. All organizations where $E_i = e$ belong to treatment cohort $e \in E$.

To construct comparable event-time paths for control organizations, I assign each control organization a *quasi-treatment date* Q_i, drawn from the empirical distribution of treatment dates among treated units, conditional on the control unit's first appearance in the sample. For treated organizations, I set $Q_i = E_i$.

The concept of a quasi-treatment date allows me to assign control organizations reference periods. Event time is defined relative to the quasi-treatment date as $k = t - Q_i$, where $k < 0$ denotes pre-periods, and $k > 0$ denotes post-periods.[23]

I use three organizational outcomes to measure resilience at time t: the number of pull requests opened, the number of pull requests merged, and the number of new software releases. Let $Y_{i,t}$ denote a generic outcome. To ensure cross-organization comparability, I standardize each organization's outcome values using the organization-level mean and standard deviation from its five most recent pre-periods ($-5 \leq k \leq -1$), yielding standardized outcomes $Y_{i,t}^{SD}$ measured in within-organization standard deviations.

I impose two additional sample restrictions. First, all organizations must have exactly one key member at the quasi-treatment date, ensuring that differences between treated and control organizations are not driven by differences in the organizational structure. Second, each organization must have observed activity for at least five pre- and five post-periods. The final sample features 681 OSS organizations and includes major GitHub projects, such as `pytorch/pytorch`, the flagship

---

[23]I abuse notation slightly not specifying that event time k_i is organization-specific to maintain simplicity.

Figure 2: Descriptive Sample Statistics

Panel (a): Outcome summary statistics

Panel (b): Distribution of Dates

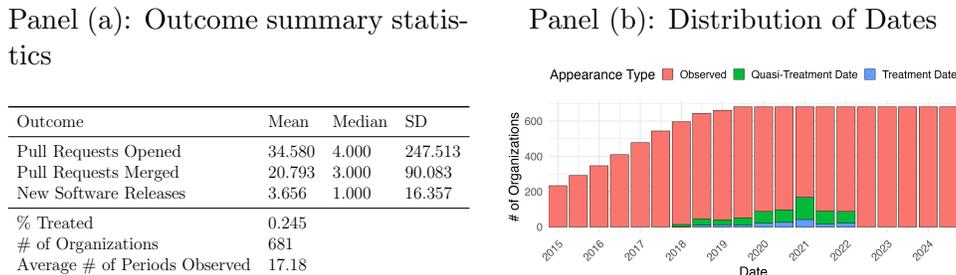| Outcome	Mean	Median	SD
Pull Requests Opened	34.580	4.000	247.513
Pull Requests Merged	20.793	3.000	90.083
New Software Releases	3.656	1.000	16.357
% Treated	0.245		
# of Organizations	681		
Average # of Periods Observed	17.18		

Figure notes: Panel 2a provides summary statistics for the unstandardized outcomes over all observed periods. The % treated and average periods observed statistics are organization-level averages. Panel 2b depicts the count distribution of observed dates, quasi-treament dates and treatment dates across all organizations.

deep learning framework; `googleapis/google-cloud-Python`, the Python Google Cloud software development kit (SDK); and `boto/boto3`, the Python Amazon Web Services SDK. Figure 2 provides summary statistics for the outcomes and project distribution, and plots the distribution of observed, quasi-treatment and treatment dates.

3 Organizational Practices

I investigate the relationship between five organizational practices—*collaboration*, *knowledge level*, *discussion quality*, *investment in new talent*, and *problem-solving routines*—and organizational resilience. The choice of these practices is motivated by prior work in management and software engineering, as discussed in the Related Literature section. Each practice includes several measures. For each organization, I compute each measure over its five most recent pre-periods and standardize the resulting values. I then aggregate the standardized measures within each practice using the first principal component to form a composite "practice score." Higher practice scores correspond to greater adoption of the practice within the organization. The following sections describe the measures within each practice and the construction of their composite scores.

3.1 Collaboration

Collaboration can enhance resilience to member departures because it ensures multiple people have experience with solving a problem. This mitigates the risk the ability to complete a task may be retained when someone leaves. I measure collaboration in an organization using five measures: the average number of participants per discussion, the percentage of discussions involving multiple members, and the degree to which participation is concentrated among a few individuals or dispersed across many—measured using the Herfindahl–Hirschman Index (HHI) for all discussions, issue threads, and pull request threads respectively. The first principal component (PC) is statistically meaningful, explaining 53.8% of the total variance across the five measures.

Table 1 provides detailed definitions of all 5 measures and each measure's weights in the first PC. Organizations with higher collaboration scores are more collaborative. For instance, a higher average number of participants per discussion increases the collaboration score, whereas organizations with more discussions that are concentrated among very few individuals (i.e., those with a higher HHI value) reduce it.

3.2   Knowledge Level

Organizations with higher knowledge levels are more resilient because each member has the knowledge to handle many tasks. This means that the loss of one member does not necessarily cause the organization loses all the knowledge held by that member. I measure an organization's knowledge level using two measures: the % of members involved in both issue and pull request threads; and the average # of unique activities members engage in. The first PC is statistically meaningful, explaining 91.6% of the total variance across the two measures.

Table 2 provides detailed definitions of both measures. Organizations with higher knowledge level scores have more knowledgeable members, which suggests higher levels of knowledge level in the organization. For example, organizations with a higher proportion of members working on both issues and pull requests have a high knowledge level score.

3.3   Discussion Quality

I use discussion quality as an umbrella term to capture the overall sentiment of discussions and whether members are responsive to each other. Higher overall sentiment can enhance organizational resilience by fostering a more positive environment, which improves member satisfaction and increases engagement. Discussions where members are more responsive demonstrate increased commitment to problem-solving, which can also improve member satisfaction. I measure discussion quality using five measures: the percentage of discussion threads that receive responses; the average number of days required for a response; and the average overall, positive, and negative sentiment scores, respectively, of all discussions, computed using the VADER algorithm (Hutto and Gilbert, 2014). The first PC component is statistically meaningful, explaining 47.7% of the total variance across the five measures.

Table 3 provides detailed definitions of all five measures. Higher discussion quality scores indicate that organizations have more positive discussions and address issues in a more timely manner. Organizations with shorter average response times or more positive sentiment tend to have higher discussion-quality scores.

3.4   Investment in New Talent

Organizations can also enhance resilience by establishing processes that facilitate the onboarding of new members. Such processes may be particularly valuable post-departure, when experienced members have limited capacity to onboard newcomers because the key member's departure means they have to assume additional responsibilities. I measure organizational investment in new talent using three measures: whether organizations have tasks labeled as "good first issues" that are appropriately difficult for new members to tackle; a contributing guide that tells new members how they can contribute to the organization; and an organizational code of conduct. The first PC is statistically meaningful, explaining 51.6% of the total variance across the three measures.

Table 4 provides detailed definitions of all three measures. Higher investment in new talent scores indicate that organizations have more processes to onboard and integrate new members automatically. For example, organizations with "good first issues" and a contributing guide have higher investment in new talent scores than those who do not.

3.5 Problem-Solving Routines

Problem-solving routines can enhance organizational resilience by establishing structured plans that guide action during times of crisis. I measure whether organizations have problem-solving routines using six measures: how frequently tasks are labeled, whether members are explicitly assigned to tasks, whether members are explicitly assigned to review tasks, whether the organization uses templates to systematize issues and pull requests, and whether it formally designates individuals as responsible for specific parts of the codebase. The first PC is statistically meaningful, explaining 35.9% of the total variance across the six measures.

Table 5 provides detailed definitions of all six measures. Higher problem-solving routine scores indicate that organizations have adopted more routines to systematize problem solving. For example, organizations that assign tags or use issue templates have higher problem-solving routine scores than those who do not.

4 Estimating Organizational Resilience

Section 4.1 estimates average organizational resilience across all organizations following the departure of a key member. I present novel evidence that key member departures have large, persistent negative effects on software development activity in open source software organizations. Section 4.2 uses a simple procedure, the *Median Split*, to test the hypotheses that average resilience increases in each organizational practice score. I find that higher scores do not suggest increased resilience to departure. Section 4.3 discusses the Median Split's limitations for hypothesis discovery.

4.1 Estimating Average Organizational Resilience

Average organizational resilience is the average causal effect of a key member departure on organizational outcomes. Because departure dates vary across organizations, I follow Sun and Abraham (2021) and estimate:

$$Y_{i,t}^{SD} = \alpha_i + \lambda_t + \sum_{e \neq \infty} \sum_{k \neq -1} \delta_{e,k} Z_{i,t}^{e,k} + \epsilon_{i,t}, \tag{1}$$

where $Z_{i,t}^{e,k} = 1$ when organization i belongs to treatment cohort e and is at event time k. Outcomes are the standardized counts of pull requests opened, pull requests merged, or new software releases. I include organization and time fixed effects, and cluster standard errors by organization.

The estimand of interest is average organizational resilience k periods after a key member departure, defined as:

$$\hat{\delta}_k = \sum_e \hat{\delta}_{e,k} \Pr(E_i = e \mid \text{treatment cohort } e \text{ is observed at event time } k) \tag{2}$$

Under two identifying assumptions, $\hat{\delta}_k$ is an unbiased and consistent estimator of average organizational resilience k periods after a key member's departure.[24]

Assumption 1 (Parallel Trends). *For all time periods $s \neq t$,*

$$\mathbb{E}[Y_{i,s}^{SD,\infty} - Y_{i,t}^{SD,\infty} \mid E_i = e] \quad \text{is the same across all treatment cohorts } e$$

---

[24]Since the sample shares of each treatment cohort are unbiased and consistent estimators for the population share, $\hat{\delta}_k$ will also be unbiased and consistent.

Figure 3: Average Organizational Resilience

Panel (a): Pull requests opened   Panel (b): Pull requests merged   Panel (c): New software releases
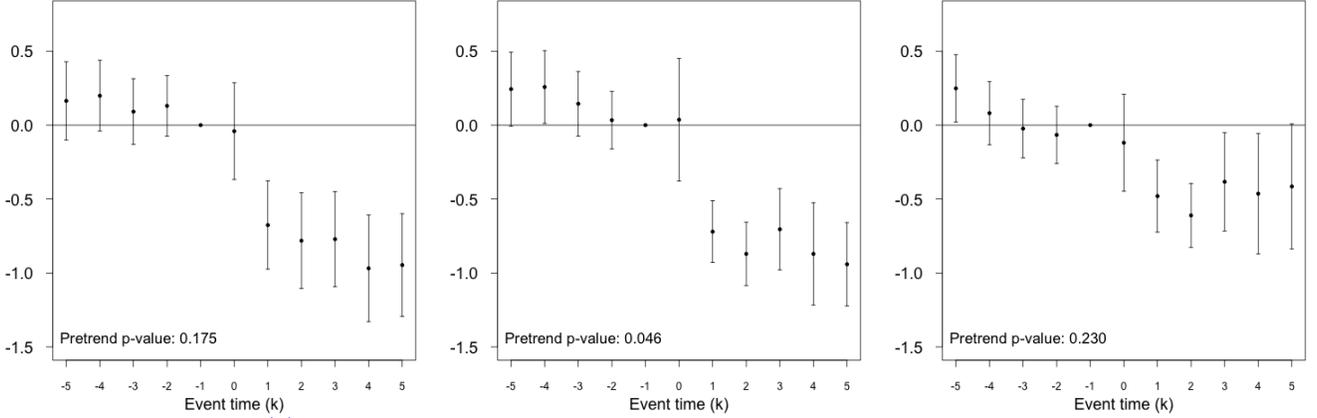

Figure notes: Panel (a) presents event study estimates of the impact of key member departures on the number of pull requests opened. Outcomes are standardized using each organization's mean and standard deviation over the five pre-periods, and each period spans six months. Estimates, expressed in standard-deviation units, are obtained using Sun and Abraham (2021). Confidence intervals are 95% intervals based on asymptotic standard errors clustered at the organization level. The reported pre-trend p-value is from a Wald test of the null hypothesis that all pre-period coefficients equal zero. Panels (b) and (c) repeat this analysis using the number of pull requests merged and the number of new software releases as outcomes.

$Y_{i,s}^{SD,\infty}$ is the potential outcome for organization i if it were never treated. The parallel trends assumption is plausible because as discussed in Section 2.4, abrupt key member departures are driven by major life changes unrelated to the OSS organization, not unobservable organizational trends that might affect the outcome.

Assumption 2 (No Anticipation).

$$\mathbb{E}[Y_{i,s}^{SD,\infty} - Y_{i,s}^{SD} \mid E_i = e] = 0 \quad \text{for all pre-periods } s < Q_i$$

A violation of the no anticipation assumption would require an OSS organization to alter its behavior months before a key member's departure. Because each period in my analysis spans six months and departure only occurs during Q_i, the organization would need substantial advance notice for any adjustment to violate the no anticipation assumption. This seems unlikely for most departures For example, job changes in the tech sector typically involve notice periods of only two weeks. Even in cases where departures are predictable—such as graduation—the organization would still need to start adjusting several months in advance for the assumption to be violated.

Figure 3 depicts event study estimates of Equation 2 for each of the three organizational outcomes—pull requests opened, pull requests merged, and new software releases—over 5 pre- and post-periods. To the best of my knowledge, Figure 3 presents the first causal estimates of the average impact of key member departures on organizational software development and release outcomes. The abrupt departure of key members has a significant negative effect on all organizational outcomes, consistent with evidence that departures from OSS organizations can be highly disruptive (Rashid, Clarke and O'Connor, 2019). These declines appear immediately after the departure, persist for at least five post-periods, and remain relatively stable. On average over the five post-periods, a key member's departure reduces pull requests opened by 0.83 standard deviations, pull requests merged by 0.82 standard deviations, and software releases produced by

Figure 4: Average Organizational Resilience by Collaboration Score

Panel (a): Pull requests opened   Panel (b): Pull requests merged   Panel (c): New software releases
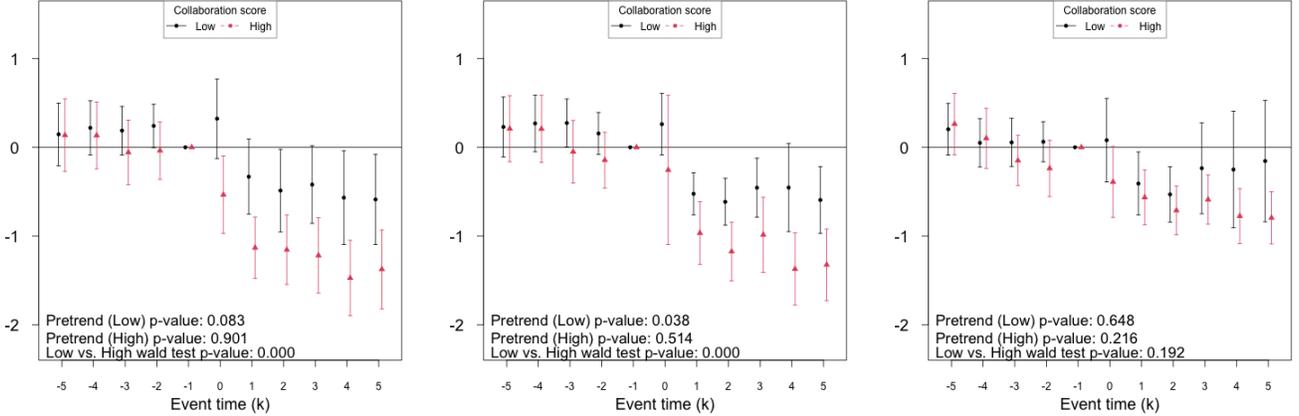

Figure notes: Panel (a) presents separate event study estimates of the impact of key member departures on the number of pull requests opened for organizations with high and low collaboration scores. Organizations are classified as having high collaboration scores if their collaboration score is above the cross-organization median, and low otherwise. Outcomes are standardized using each organization's mean and standard deviation over the five pre-periods, and each period spans six months. Estimates, expressed in standard-deviation units, are obtained using Sun and Abraham (2021). Confidence intervals are 95% intervals based on asymptotic standard errors clustered at the organization level. The two reported pre-trend p-values are Wald tests of the null that all pre-period coefficients equal zero for the high- and low-collaboration subsets. The final p-value is from a Wald test of the null that post-period event-study coefficients do not differ between the two subsets. Panels (b) and (c) repeat this analysis using the number of pull requests merged and the number of new software releases as outcomes.

0.47 standard deviations. Since the quantity of new software releases declines, users will directly experience the effects of the departure through less timely bug fixes and fewer new features. Given that reliance on OSS is widespread, this indicates that low organizational resilience in OSS projects may have substantial downstream economic consequences.

The bottom-left of each panel reports the p-value from the Wald test of the pre-trend null hypothesis $\delta_k = 0$ for the five plotted pre-periods, following Freyaldenhoven et al. 2025. For two of the three organizational outcomes, I fail to reject the null of no pre-trends. For pull requests merged, the null is marginally rejected at the 5% level. I do not consider this especially problematic for two reasons: the estimated effects are the opposite sign of the pre-trend coefficients and the pre-trend estimates closest to the treatment date are statistically indistinguishable from zero. Unobserved trends most threatening to identification would likely manifest as statistically significant pre-trends in the periods immediately prior to departure—which I do not observe.

4.2   Simple Tests of Organizational Resilience

Next, I test the hypothesis that organizations with higher organizational practice scores are more resilient. Each organization has a vector of five practice scores X_i, with X_{ij} denoting the score for practice j. For each practice j, I apply the *Median Split*, separating organizations into high and low score groups, $C_{ij} \in \{\text{high}, \text{low}\}$, depending on whether X_{ij} lies above or below the cross-organizational median.

Figure 4 depicts separate event study estimates of average organizational resilience for organizations with high and low collaboration scores, across all organizational outcomes. Let $\hat{\delta}_k^{high,collab}$ denote the estimates from Equations 1– 2 for organizations with high collaboration scores, and

$\hat{\delta}_k^{low,collab}$ the corresponding estimates for organizations with low collaboration scores. Each event study plot reports three p-values. In descending order, the first two report the p-value of the pre-trend Wald test for high and low collaboration score organizations, respectively. The last p-value is from a Wald test evaluating the null hypothesis that the event study estimates for high and low collaboration score organizations do not differ for all event times $1 \leq k \leq 5$.[25] Rejecting this null indicates that organizations with high collaboration scores exhibit different levels of organizational resilience than those with low scores post-departure.

Organizations with high collaboration scores are not more resilient after a key member's departure; if anything, they experience larger declines in pull requests opened and merged, and these differences persist across all post-periods. This pattern suggests that if high collaboration scores do increase resilience, it must do so through its interaction with other organizational practices.

Appendix Figures A1– A2 extend the analysis in Figure 4 to the remaining organizational practice categories: knowledge level, discussion quality, investment in new talent, and problem-solving routines. I find that no individual organizational practice independently predicts greater resilience as its adoption increases. This finding contrasts sharply with findings from the management and software engineering that increased adoption of these practices strengthens resilience.

Two identifying assumptions allow $\hat{\delta}_k^{high,collab}$ and $\hat{\delta}_k^{low,collab}$ (and their counterparts for other organizational practices) to be interpreted as causal effects (caused by departures). Given an organization's multi-valued vector of all five organizational practice scores X_i,

Assumption 3 (Conditional Parallel Trends). *For all time periods $s \neq t$,*

$$\mathbb{E}[Y_{i,s}^{SD,\infty} - Y_{i,t}^{SD,\infty} \mid E_i = e, X_i = x] \quad \text{is the same across all treatment cohorts } e,$$

Assumption 4 (Conditional No Anticipation).

$$\mathbb{E}[Y_{i,s}^{SD,\infty} - Y_{i,s}^{SD} \mid E_i = e, X_i = x] = 0 \quad \text{for all pre-periods } s < Q_i.$$

I consider conditional parallel trends plausible for two reasons. First, because abrupt key member departures are typically driven by major life changes unrelated to the OSS organization, conditioning on organizational practices should not induce differential correlation with unobserved trends that affect outcomes. Second, Assumption 3 holds constant organizational practices that could plausibly affect outcomes, reducing cross-organizational differences along an additional, substantively meaningful dimension. I also find the conditional anticipation assumption plausible because although better organized organizations might require or expect longer notice periods, it is still unlikely the notice periods would be as long as several months.

Assumption 3 and Assumption 4 identify $\hat{\delta}_k^{high,collab}$ as the causal estimate of how departures affect organizations with high collaboration scores (or any potentially multidimensional combination of organizational practice scores). However, without further assumptions, $\hat{\delta}_k^{high,collab}$ cannot be interpreted as the causal effect of high collaboration scores on organizational resilience to key member departures. Unless the possibility that the observed estimate of resilience is driven by other (potentially unobserved) organizational practice is ruled out, collaboration scores cannot be interpreted as a causal driver of resilience. Instead, the estimate should be interpreted as descriptive evidence of the relationship between high collaboration scores and resilience. The difficulty of assigning causality is encountered by most empirical work; short of randomly assigning organizational practices, a practice should only be viewed as causal if it is clear that the

---

[25]The p-value is from the Wald test $H_0 : \hat{\delta}_k^{high,collab} - \hat{\delta}_k^{low,collab} = 0$ for all $1 \leq k \leq 5$.

practice's mechanisms directly enhance resilience. Establishing causal drivers is difficult; even in Bloom et al. (2012), where management practice advice is randomly provisioned, organizations still endogenously choose whether to adopt the practices.

4.3 Flaws of the Median Split

Although Figure 4 and Appendix Figures A1– A2 suggest that higher organizational practice scores do not increase organizational resilience, this null result may simply reflect the limitations of the Median Split. The Median Split is poorly suited to discover how practices affect resilience because it imposes a restrictive model of their relationship that ignores interactive effects of practices and groups organizations without using data-driven splits.

Extending the Median Split is a poor recipe for discovering how multiple organizational practices interactively affect organizational resilience. Allowing for interactions is important because the empirical literature on complementarities finds that firms benefit from simultaneously adopting bundles of practices, as opposed to individual ones (Ichniowski, Shaw and Prennushi, 1997). Due to the curse of dimensionality, the number of possible interactions between organizational practices increases factorially in the number of variables in each interaction. Naively reporting covariate combinations where I observe statistically significant differences across subgroups complicates inference due to multiplicity of testing.

Second, since the Median Split does not use a data-driven approach when partitioning the space of organizational practice scores, some subsets with heterogeneity may be omitted. It is not necessarily true (and likely incorrect) to assume ex-ante that for any given organizational practice, splitting along the median and using two subsets is the best way to model heterogeneity in resilience induced by practices. For example, suppose that organizational resilience is highest for only values at the left and right tail of the score distribution for a particular organizational practice. In this case, the Wald test may report that the coefficients for event studies from high and low score subsets produced by the Median Split are statistically indistinguishable. Assuming based off this that no heterogeneity exists is incorrect; the relationship between the practice score and resilience is just nonlinear.

5   Flexibly Modeling Organizational Resilience

To address the limitations of the Median Split, I turn to machine learning methods that can flexibly model the relationship between organizational practices and organizational resilience. Building on recent advances in the heterogeneous treatment effects literature, I develop the *Generalized Random Forest Event Study* (Liao, 2025), which adapts the generalized random forest (Athey, Tibshirani and Wager, 2019) to estimate heterogeneous dynamic treatment effects in event study settings. Section 5.1 describes how this method can be flexibly applied to model the relationship between practices and resilience. Section 5.2 then shows how I use the forest's out-of-sample predictions to construct the *Forest Split*, a data-driven classification procedure that successfully separates organizations into high and low resilience subsets based on their practices.

5.1   The Generalized Random Forest Event Study

A standard approach to estimating heterogeneous treatment effects is to use the generalized random forest (Athey, Tibshirani and Wager, 2019), which estimates heterogeneous treatment effects

by adaptively splitting the covariate space to group together observations with comparable treatment effects, before averaging treatment effects within these groups. This splitting procedure is done thousands of times, each on a different random subsample of the dataset; the set of splits associated with each random subsample is called a *tree*. Modeling organizational resilience using the generalized random forest addresses both flaws discussed in Section 4.3, as the forest allows for interactions and adaptively splits the covariate space to group observations based on treatment effect similarity.

The generalized random forest cannot be applied to my setting out-of-the-box since my estimand of interest is a conditional dynamic treatment effect $\delta_k(X_i)$. To address this, I develop the *Generalized Random Forest Event Study* in a separate note (Liao, 2025). I show that estimating heterogeneous treatment effects in event study settings with staggered adoption and unit-invariant covariates is equivalent to estimating a special case of the conditional linear model with binary regressors and propensity scores tailored to the panel data and staggered treatment adoption setting. Because the conditional linear model is unbiased and consistent when estimated using generalized random forests (Athey, Tibshirani and Wager, 2019), the heterogeneous treatment effects estimated by the Generalized Random Forest Event Study inherit these properties under assumptions tailored to the panel setting.

The Generalized Random Forest Event Study models the standardized organizational outcome $Y_{i,t}^{SD}$ as

$$Y_{i,t}^{SD} = \alpha_i + f_t(X_i) + \sum_{e \neq \infty} \sum_{k \neq -1} \delta_{e,k}(X_i) Z_{i,t}^{e,k} + \varepsilon_{i,t} \tag{3}$$

Identification of $\delta_{e,k}(X_i)$ follows based on conditional parallel trends (Assumption 3) and conditional no anticipation (Assumption 4). I also assume that the regularity conditions in Section 3 of Athey, Tibshirani and Wager (2019) hold.

In addition to estimating heterogeneous treatment effects, Equation 3 replaces the time fixed effects from Equation 1 with flexible time-varying functions of covariates $f_t(X_i)$. $f_t(X_i)$ captures baseline heterogeneity induced by covariates across time. Organizational practices affect resilience, and hence, organizational outcomes. Hence, allowing for organizational practice-dependent baseline heterogeneity is important because there is no fundamental reason why practices would affect organizational outcomes only during the treatment period.

The estimation procedure for $\hat{\delta}_{e,k}(x)$ follows Athey, Tibshirani and Wager (2019), with two caveats.[26] First, as discussed in Appendix C of Liao (2025), treatment propensity scores are adapted to the panel data and staggered treatment adoption setting.[27] Second, any given organization i, $\hat{\delta}_{e,k}(X_i)$ is estimated following a k-fold procedure.[28] Using a $k-$fold procedure ensures that small biases from the outcome and propensity score models used for estimation will not affect $\hat{\delta}_{e,k}(x)$.

For each organization's cohort-event time estimates $\hat{\delta}_{e,k}(X_i)$, I calculate the doubly robust equivalent $\hat{\psi}_k(X_i)$, using the outcome model and treatment propensity function from the trained forest.[29] For an organization i with quasi-treatment cohort e, the doubly robust average post-

---

[26]I use the default hyperparameters as of version 2.4.0.

[27]Athey, Tibshirani and Wager (2019) use a probability forest to estimate treatment propensity.

[28]Instead of a k-fold estimation procedure, Athey, Tibshirani and Wager (2019) use out-of-bag predictions, which for a given observation i, exclude trees trained on subsamples that include i. $k-$fold estimation is more time-intensive but is precisely covered by theory.

[29]As described in Appendix D of Liao (2025), I adapt the procedure from Uysal (2015), which implements propensity scores estimation when treatments are multivalued, to my staggered treatment adoption setting.

17

Algorithm 1 Forest Split: k-Fold Procedure

1: Partition organizations into k folds $S_P = \{P_1, \ldots, P_k\}$
2: **for** each fold $P_j \in S_P$ **do**   ▷ Cross-fitting loop
3:      Let the $k-1$ remaining folds be denoted $P_{-j} = S_P \setminus P_j$
4:      Estimate Equation 3 using organizations $i \in P_{-j}$ and covariates $\mathbf{F}_i$ to obtain $\{\hat{\delta}_{e,k}^{P_{-j}}(\cdot)\}_{e \in E, -5 \leq k \leq 5}$
5:      For each organization $i \in P_j$, predict post-period effects $\{\hat{\delta}_{e,k}^{P_{-j}}(\mathbf{F}_i)\}_{e \in E, -5 \leq k \leq 5}$
6: **end for**
7: For each organization i, compute the doubly robust average post-period treatment effect $\hat{\psi}(\mathbf{F}_i)$ following Equation 4
8: Let $m \leftarrow \text{median}\{\hat{\psi}(\mathbf{F}_i)\}$
9: **for** each organization i **do**
10:     **if** $\hat{\psi}(\mathbf{F}_i) < m$ **then**
11:         Assign i to the low resilience subset
12:     **else**
13:         Assign i to the high resilience subset
14:     **end if**
15: **end for**

---

period treatment effect is defined as:

$$\hat{\psi}(X_i) = \frac{1}{5} \sum_{k>0}^{5} \hat{\psi}_{e,k}(X_i) \tag{4}$$

5.2 The Forest Split

To assess whether organizational practices affect organizational resilience, I define the *Forest Split*, which categorizes organizations as high (low) resilience if their doubly robust post-period average treatment effect $\hat{\psi}(X_i)$ is above (below) the organization-wide median. Algorithm 1 describes each step of the Forest Split, using a $k-$fold estimation procedure.

As discussed in Section 4.3, a limitation of the Median Split is that it imposes a restrictive model of how organizational practices affect resilience by ruling out interactions across practices. The Generalized Random Forest Event Study relaxes this restriction, as forests naturally capture interactions among covariates when they are statistically informative. A natural question, then, is whether this more flexible model can better distinguish between high and low resilience organizations. I evaluate this by applying Algorithm 1 to assign organizations to high and low resilience subsets. In this implementation, Equation 3 is estimated using organizational practice scores coarsened by the Median Split, C_i, as covariate inputs and the standardized pull requests merged outcome, $Y_{i,t}^{SD}$, as the outcome.

Figure 5 depicts separate event study estimates of average organizational resilience for organizations classified as high and low resilience, across all organizational outcomes. I find that the difference in average organizational resilience between the high and low resilience subsets is both statistically significant and economically large. On average across the five post-periods, organizations in the high resilience subset open 1.80 standard deviations more pull requests, merge 1.99 more pull requests, and produce 0.81 standard deviations more software releases than those in the

Figure 5: Average Organizational Resilience Using the Forest Split and Coarsened Practice Scores

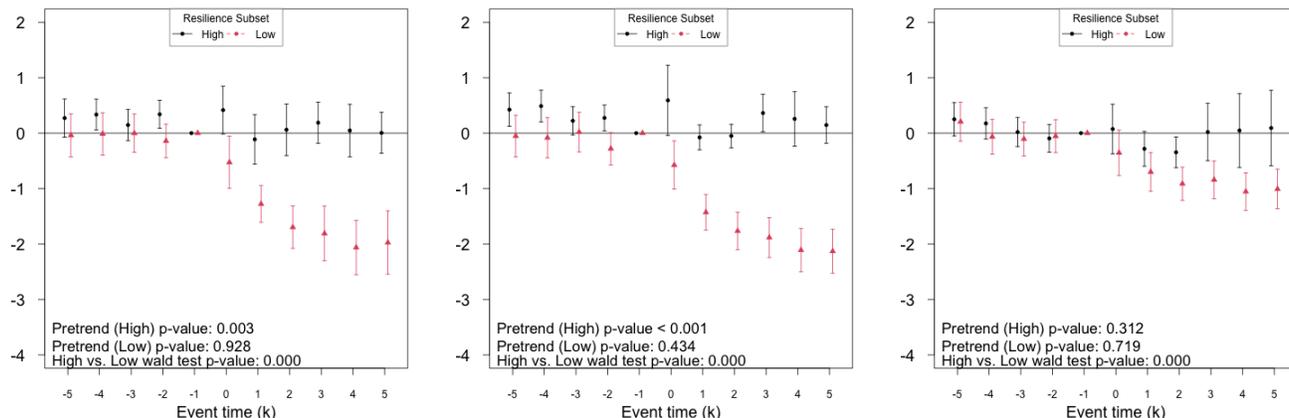Panel (a): Pull requests opened Panel (b): Pull requests merged Panel (c): New software releases

Figure notes: Panel (a) presents event study estimates of the impact of key member departures on the number of pull requests opened for organizations in the high and low resilience subset. Organizations are assigned to the high resilience subset when their estimated doubly robust average post-period treatment effect–computed using organizational practice scores coarsened into binary high/low values–exceeds the cross-organizational median. Organizations with below-median effects belong to the low resilience subset. Outcomes are standardized using each organization's mean and standard deviation over the five pre-periods, and each period spans six months. Estimates, expressed in standard-deviation units, are obtained using Sun and Abraham (2021). Confidence intervals are 95% intervals based on asymptotic standard errors clustered at the organization level. The two reported pre-trend p-values are Wald tests of the null that all pre-period coefficients equal zero for the high- and low-collaboration subsets. The final p-value is from a Wald test of the null that post-period event-study coefficients do not differ between the two subsets. Panels (b) and (c) repeat this analysis using the number of pull requests merged and the number of new software releases as outcomes.

low resilience subset. Because the heterogeneous treatment effects used to separate organizations are out-of-sample predictions, the forest's explanatory power likely reflects its ability to capture genuine statistical relationships between practices and resilience rather than overfitting. That binary coarsenings of these practices are able to predict such large differences in resilience is also striking. Moreover, because the forest used only the pull requests merged outcome to predict and separate organizations into high and low resilience subsets, the same resilience classification is used to show results for the other outcomes in Figure 5.

This indicates that the same combinations of organizational practices underpin resilience across all three outcomes.

The second limitation of the Median Split is that it partitions organizations along a naive threshold rather than using data-driven splits of organizational practices. To assess the value of allowing the model to partition organizations along a finer score space, I apply Algorithm 1 using the continuous organizational practice scores X_i as covariates to assign organizations to high and low resilience subsets. Figure 6 plots separate event study estimates for the high and low resilience subsets defined by this procedure across all organizational outcomes. I find that the differences between high and low resilience organizations grows even larger when the model can partition along the full continuous distribution of practice scores. On average across the five post-periods, organizations in the high resilience subset open 1.86 standard deviations more pull requests, merge 2.16 more pull requests, and produce 1.36 standard deviations more software releases than organizations predicted to be less resilient. However, the fact that the largest gains in predictive power arise from allowing multiple practices to interact and jointly affect resilience

Figure 6: Average Organizational Resilience Using the Forest Split and Continuous Practice Scores

Panel (a): Pull requests opened   Panel (b): Pull requests merged   Panel (c): New software releases
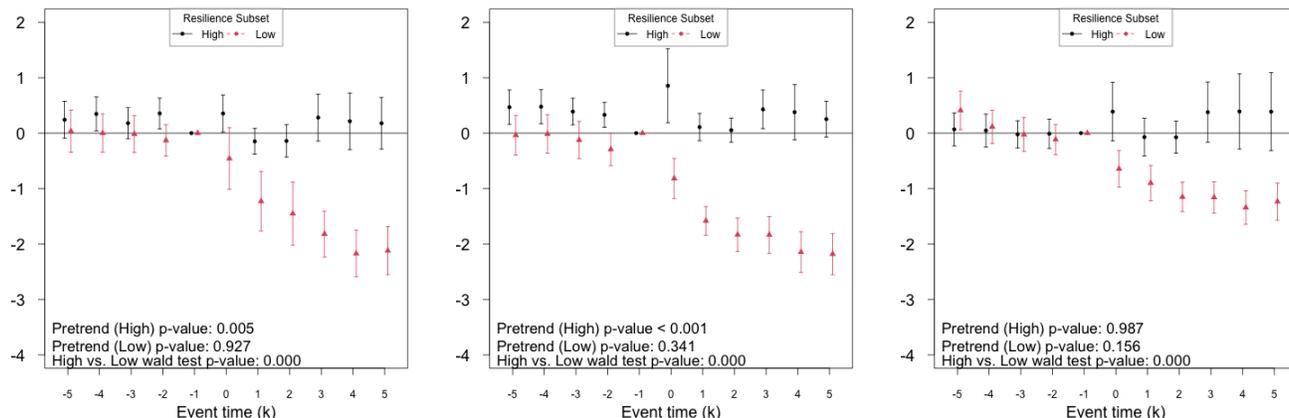

Figure notes: Panel (a) presents event study estimates of the impact of key member departures on the number of pull requests opened for organizations in the high and low resilience subset. Organizations are assigned to the high resilience subset when their estimated doubly robust average post-period treatment effect–computed using continuous organizational practice scores–exceeds the cross-organizational median. Organizations with below-median effects belong to the low resilience subset. Outcomes are standardized using each organization's mean and standard deviation over the five pre-periods, and each period spans six months. Estimates, expressed in standard-deviation units, are obtained using Sun and Abraham (2021). Confidence intervals are 95% intervals based on asymptotic standard errors clustered at the organization level. The two reported pre-trend p-values are Wald tests of the null that all pre-period coefficients equal zero for the high- and low-collaboration subsets. The final p-value is from a Wald test of the null that post-period event-study coefficients do not differ between the two subsets. Panels (b) and (c) repeat this analysis using the number of pull requests merged and the number of new software releases as outcomes.

suggests that data-driven splits are a second-order modeling concern.

One concern with Figure 5 and Figure 6 is that in the panels where pull request activity is used as an outcome, I observe statistically significant pre-trends for highly resilient organizations. These pre-trends are concerning because they indicate that, prior to the departure, organizations in the high resilience subset were already more productive (relative to their reference period) compared to their never-treated counterparts. Consequently, average resilience estimates for these organizations are likely biased upward, as they capture not only resilience to the departure but also the underlying upward trajectory in outcomes evident in Panels 5a– 5b and Panels 6a– 6b. Although the pre-trend magnitudes are substantially smaller than the difference in resilience across high and low resilience organizations and do not follow any perceptible upward or downward trend, future work should consider adopting methods from Rambachan and Roth (2023) to perform robust inference under potential violations of the parallel trends assumption.

A significant methodological contribution of the Generalized Random Forest Event Study is that it models treatment *dynamics*. One natural question is whether organizational resilience is persistent–that is, are the organizations that are most resilient immediately after a departure also those that remain resilient in the long run? To test this hypothesis, I define an early and a late resilience classification that describes whether organizations were resilient immediately versus several periods following departure. I do so by adapting the metric used in Algorithm 1 to subset organizations into high and low resilience subsets. To separate organizations into the early high versus early low resilience subsets, I evaluate whether an organization's doubly robust first post-

Figure 7: Average Organizational Resilience Using the Forest Split and Continuous Practice Scores

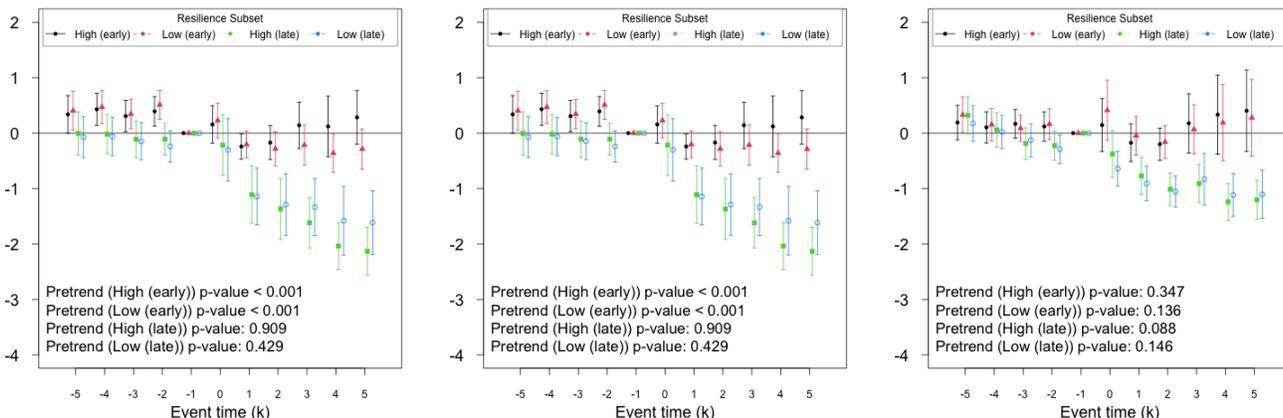Panel (a): Pull requests opened　　Panel (b): Pull requests merged　　Panel (c): New software releases

Figure notes: Panel (a) presents event study estimates of the impact of key member departures on the number of pull requests opened for organizations in the early high, early low, late high and low low resilience subset. Organizations are assigned to the early (late) high resilience subset when their estimated doubly robust first (fifth) post-period treatment effect–computed using continuous organizational practice scores–exceeds the cross-organizational median. Organizations with below-median effects belong to the early (late) low resilience subset. Outcomes are standardized using each organization's mean and standard deviation over the five pre-periods, and each period spans six months. Estimates, expressed in standard-deviation units, are obtained using Sun and Abraham (2021). Confidence intervals are 95% intervals based on asymptotic standard errors clustered at the organization level. The two reported pre-trend p-values are Wald tests of the null that all pre-period coefficients equal zero for the high- and low-collaboration subsets. The final p-value is from a Wald test of the null that post-period event-study coefficients do not differ between the two subsets. Panels (b) and (c) repeat this analysis using the number of pull requests merged and the number of new software releases as outcomes.

period treatment effect $\hat{\psi}_1(X_i)$[30] is above or below the across-organization median. Analogously, organizations are separated into their late low and high resilience subsets using the doubly robust fifth post-period treatment effect $\hat{\psi}_5(X_i)$.

Figure 7 depicts separate event study estimates of average organizational resilience for organizations belonging to the early low, early high, late low and late high resilience subset. First, irrespective of whether organizations belong to the early high or late high resilience subset, they fare substantially better than their low resilience counterparts post-departure. Second, organizations in the early high resilience subset tend to be slightly less resilient several periods following departure than organizations who are most resilient in the final period post-departure. This suggests that practices that support short-term resilience also support longer-term resilience, although they may not be the optimal set for maximizing long-term resilience.

Figures 5–7 present several takeaways. First, the true relationship between the five organizational practices and resilience involves multiple practices that jointly and interactively affect resilience. Although adopting any individual practice cannot enhance an organization's resilience, there are combinations of practices that are associated with enhanced resilience. Second, I document substantial variation in average organizational resilience across OSS organizations, demonstrate that these differences are large and persistent, and show that the organizations exhibiting high resilience on average represent a meaningful fraction of the population. What remains unclear is which combinations of organizational practices are systematically associated with higher

---

[30]I simplify notation by defining, for an organization i with quasi-treatment period e, the k-th doubly robust post-period treatment effect $\hat{\psi}_k(X_i) = \hat{\psi}_{e,k}(X_i)$

resilience, and whether these combinations include higher practice scores. Answering these questions is the focus of the next section.

6   Interpreting the Random Forest

The Forest Split successfully separates organizations into high and low organizational resilience subsets using just 5 organizational practice scores. Importantly, the Forest Split's predictions are out of sample, which suggests that the discovered relationships about how practices affect resilience are generalizable. Yet, as with many machine-learning methods, the relationships used by the forest to generate resilience subsets remain a black box. This section aims to unpack that black box to discover hypotheses about the relationship between practices and resilience.

Section 6.1 describes the characteristics of the trained forest. Finally, Section 6.2 coarsens the space of organizational practice scores to examine how resilience varies across all possible practice combinations.

6.1   Characteristics of the Random Forest

Figure 8 describes characteristics of the Generalized Random Forest Event Study that models $\delta_{e,k}(\cdot)$ from Equation 3. The described forest is trained on the full sample of organizations and the continuous organizational practice scores X_i.

Panel 8a of Figure 8 describes the distribution of tree depth among all 2000 trees in the trained forest. Over half of the trees have depth three or greater, suggesting that complex interactions between organizational practices play an important role in predicting resilience, since the forest relies on multi-step splits to separate organizations.

Panel 8b of Figure 8 describes how often the forest uses each organizational practice score to partition organizations into subsets with different treatment effects. Most organizational practices are used frequently by the forest, except for the investment in new talent score. This suggests that the forest models resilience as a function of several practices that interact.

Notice that the forest rarely splits on the investment in new talent score. There are two complementary explanations for why the forest rarely uses the investment in new talent score to partition organizations. First, it may simply play a limited role in determining resilience. Second, as Panel 8c of Figure 8 shows, the investment in new talent score is highly correlated with the problem-solving routines score, so the effect may be masked by the problem-solving routines score.

Figure 8c also shows that the other organizational practice scores are not strongly correlated. Although the absence of strong cross-practice correlations does not mean that my analysis will isolate the causal effect of each practice, it does suggest that I am likely not confounding effects associated with one practice, such as collaboration, with those of another practice, such as problem-solving routines.

6.2   Exploring the Coarsened Space of Organizational Practices

A key insight from Figure 5 is that coarsening continuous practice scores into binary high/low scores using a median split only slightly reduces the generalized random forest's ability to separate high resilience and low resilience organizations. This motivates a complementary interpretability approach: instead of working with continuous scores, I classify organizations into one of 32 categories corresponding to all possible combinations of the five binary practice indicators. Although
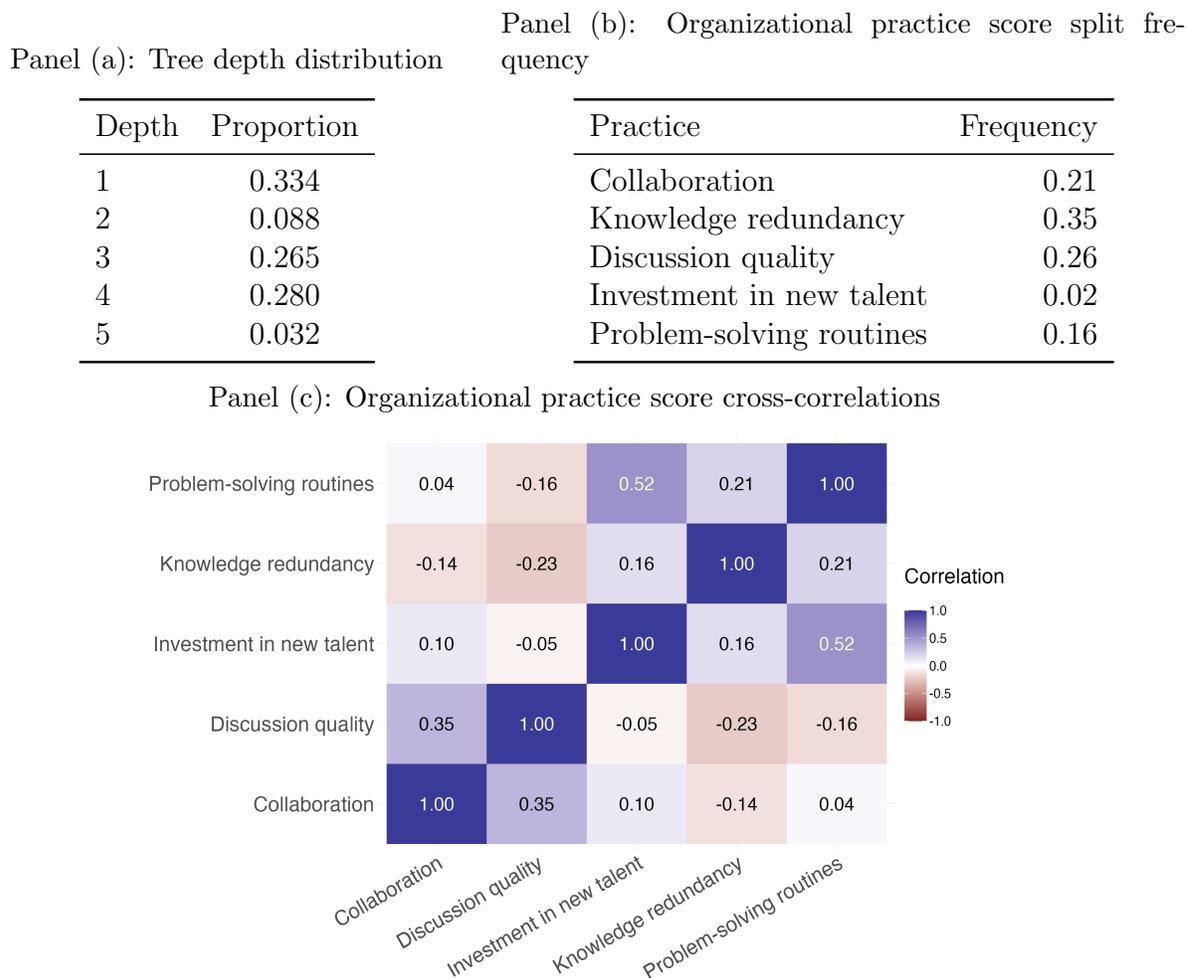
Figure 8: Characteristics of the Trained Forest

Panel (a): Tree depth distribution

Panel (b): Organizational practice score split frequency

Depth	Proportion
1	0.334
2	0.088
3	0.265
4	0.280
5	0.032

Practice	Frequency
Collaboration	0.21
Knowledge redundancy	0.35
Discussion quality	0.26
Investment in new talent	0.02
Problem-solving routines	0.16

Panel (c): Organizational practice score cross-correlations

Figure notes: Panel (a) reports the distribution of tree depth across all 2000 trees in a forest trained on all observations and the 5 continuous organizational practice scores. Panel (b) reports the frequency with which each practice score is used to split the sample, weighted by tree depth following the variable-importance measure in GRF 2.4.0. Panel (c) depicts a heatmap of correlations across all practice scores.

the coarsening discards information, I believe the modest loss in predictive precision is outweighed by the interpretability gained from reducing the continuous practice space to a manageable set of 32 combinations.

Figure 9 displays all combinations of coarsened organizational practice scores, ordered in descending order by the mean of the doubly robust average post-period treatment effect across all organizations in the combination. Recall from Section 4.2 that high collaboration scores do not suggest increased resilience, and for some outcomes, high collaboration scores suggest decreased resilience. This pattern is observed in Figure 9 as well; 9 of the 16 combinations with high collaboration scores belong in the top half of combinations, ordered by resilience, whereas the remaining 7 belong in the bottom half. Across all organizational practice scores, I again observe that high values for a practice score are not systematically associated with increased resilience.

However, high collaboration does suggest increased resilience when accompanied by high knowledge level scores within the organization and/or high discussion quality. Six of the eight combinations that contain both high collaboration and knowledge level scores belong in the top half of combinations, ordered by resilience. One explanation from the literature is that collaboration
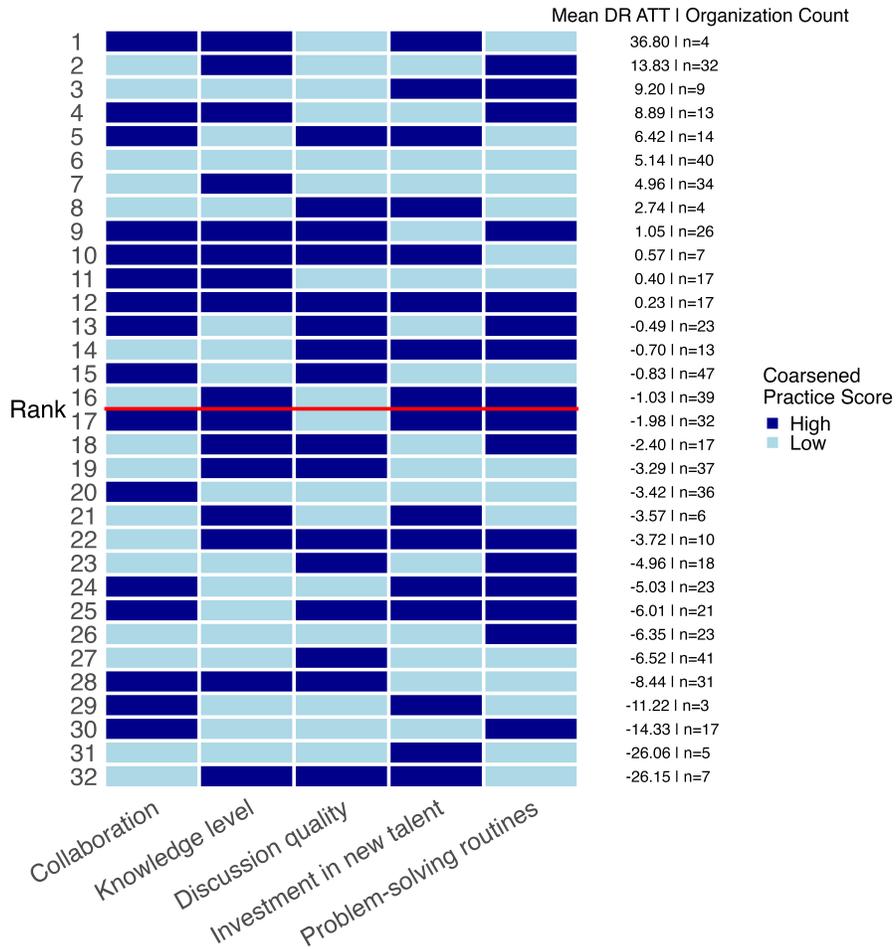
23

Figure 9: Coarsened Practice Combinations

Figure notes: Each row represents one of the 32 possible combinations of coarsened organizational practice scores, ordered from highest to lowest based on the mean doubly robust average post-period treatment effect across all organizations with that combination of practices. Each column corresponds to one organizational practice. Dark shading indicates that the practice score in that combination is high, while light shading indicates that it is low. To the right of each row, I report both the mean doubly robust average post-period treatment effect and the number of organizations with that combination of practices.

enhances resilience only when there is sufficient knowledge to transfer and share between organization members (Gressgård and Hansen, 2015). A second explanation, given that I measure practices over five periods, is that collaboration without evidence of knowledge development (proxied by high knowledge level scores) is unlikely to be effective (Fiol and Lyles, 1985).

Six of the eight practice combinations that contain both high collaboration and discussion quality scores also belong in the top half of combinations, ordered by resilience. One explanation is that collaboration is only effective at inducing resilience when coordinated well. Collaborators that are negative or slow to respond may hinder rather than enhance their team's performance, especially following disruptive situations (Hoffer Gittell, 2002). This finding also builds on existing research emphasizing that effective collaboration depends on overcoming coordination and communication challenges during "normal times" (Cross, 2020; Dietrichson, Gudmundsson and Jochem, 2022). Organizations with higher knowledge and discussion-quality scores are not systematically more resilient; rather, high collaboration scores appears to be the necessary ingredient that associates these practices with enhanced organizational resilience.

A third observation is that organizations with high collaboration, high knowledge levels, and high discussion quality scores are more resilient only when they also have either high scores in investment in new talent and/or problem-solving routines. When many knowledgeable members collaborate intensively, organizations benefit from routines that structure information exchange and from onboarding practices that prevent new members from becoming overwhelmed. Widely adopted routines help organize information and maintain alignment across the organization (Home III and Orr, 1997). In my setting, this could strengthen resilience by ensuring that other organization members are up to speed and can continue the departed member's work. Problem-solving routines and structures for investment in new talent are costly to adopt and maintain; my results suggest that they help resilience best when organizations face substantial coordination demands.

These findings show that the organizational practices emphasized in prior research do matter for resilience, but their effects depend on what other practices are adopted by the organization. High collaboration scores are associated with higher resilience only when paired with either high knowledge levels or high discussion quality, suggesting that organizations must possess the capacity to share knowledge effectively and coordinate communication for collaboration to actually support resilience. Likewise, investment in new talent and structured problem-solving routines are complementary to settings where many knowledgeable members are actively collaborating and generating substantial information flows. These are contexts in which organizational alignment requires systematization and onboarding processes most, making it easier to maintain continuity post-departure. Associated resilience benefits from high organizational practice scores arise from complementarities between having high scores across multiple practices.

7   Conclusion

In this paper, I develop measures of organizational resilience and practices using detailed activity-level data on open source software projects. I show, via a novel method combining event study and heterogeneous treatment effect estimation, that differences in organizational practices explain substantial variation in organizational resilience to key member departures. I find that although organizations with high collaboration scores are not more resilient, collaboration is a practice that benefits tremendously from complementarities. For example, organizations with high collaboration scores are more resilient when their knowledge level across members is high, or when their discussion quality scores are high.

There are several promising directions for future research. First, studying how the composition of software releases changes following key member departures can improve our understanding of the importance of OSS organizational resilience to downstream software users. To better understand the causal drivers of resilience, it would be valuable to examine the mechanisms through which organizations maintain resilience and assess whether these mechanisms align with the hypotheses discovered in this paper. Finally, since adopting organizational practices involves meaningful tradeoffs, incorporating these tradeoffs into the analytical framework would deepen our understanding of the broader relationship between practices and organizational outcomes.

Table 1: Collaboration Measures

Name	Definition	PC1 Weight
Average discussants	Average # of discussants per issue or pull request thread	0.542
Multiple discussants	% of issue or pull request threads with multiple discussants	0.400
Comment concentration	Organization-level Herfindahl–Hirschman Index based on each member's share of issue or pull request comments	-0.431
Issue thread concentration	Weighted average of thread-level Herfindahl–Hirschman Indices based on members' shares of comments within issue threads, weighted by each thread's total comment count	-0.441
Pull request comment concentration	Weighted average of thread-level Herfindahl–Hirschman Indices based on members' shares of pull request comments within pull request threads, weighted by each thread's total comment count	-0.407

Table 2: Knowledge Level Measures

Name	Definition	PC1 Weight
Both issue and pull request share	% of members involved in both issue and pull request threads	0.707
Unique activity types	Average number of unique activities each member engages in, across the following nine options: opening, commenting on, or closing issues; and opening, commenting on, leaving review comments, reviewing, or merging pull requests	0.707

Table 3: Talent Investment Measures

Name	Definition	PC1 Weight
Response rate	Percentage of discussion threads that received a response, excluding pull request reviews	0.343
Response time	Average # of days per response	-0.065
Response positivity	Average positivity score of all discussion, according to VADER (Hutto and Gilbert 2014)	0.612
Response negativity	Average negative score of all discussion, according to VADER (Hutto and Gilbert 2014)	0.533
Overall sentiment	Compound sentiment score of all discussion, according to VADER (Hutto and Gilbert 2014)	-0.469

Table 4: Investment in New Talent Measures

Name	Definition	PC1 Weight
Has Good First Issues	Whether an organization has dedicated problems specifically for new contributors to work on	0.495
Has Contributing Guide	Whether an organization has a dedicated guide to onboarding potential new members	0.622
Has Code of Conduct	Whether an organization has a dedicated guide to "define community standards, signal a welcoming and inclusive project, and outline procedures for handling abuse" (Github, 2025a)	0.607

Table 5: Problem Solving Routine Metrics

Name	Definition	PC1 Weight
Assigned Reviewers	Average % of pull requests with assigned reviewers	0.375
Assigned Assignee	Average % of issues or pull requests with a member assigned to work on it	0.422
Assigned Tag	Average % of issues or pull requests with tags labeling them	0.396
Has Codeowners	Whether an organization has a CODEOWNERS file on GitHub defining which members are in charge of the codebase	0.353
Has Issue Template	Whether an organization uses issue templates to systematize inquiries and discussion	0.478
Has Pull Request Template	Whether an organization uses pull request templates to systematize proposed codebase changes	0.414

References

Annarelli, Alessandro, and Fabio Nonino. 2016. "Strategic and Operational Management of Organizational Resilience: Current State of Research and Future Directions." *Omega*, 62: 1–18.

Athey, Susan, Julie Tibshirani, and Stefan Wager. 2019. "Generalized Random Forests." *The Annals of Statistics*, 47(2): 1148–1178.

Atkinson, Mariam Krikorian. 2023. "Organizational Resilience and Change at UMass Memorial." Accessed August 2025.

Azoulay, Pierre, Christian Fons-Rosen, and Joshua S. Graff Zivin. 2019. "Does Science Advance One Funeral at a Time?" *American Economic Review*, 109(8): 2889–2920.

Azoulay, Pierre, Joshua S. Graff Zivin, and Jialan Wang. 2010. "Superstar Extinction." *The Quarterly Journal of Economics*, 125(2): 549–589.

Bals, Fred. 2025. "Six Takeaways From the 2025 "Open Source Security and Risk Analysis" Report." Accessed February 2025.

Bao, Lingfeng, Xin Xia, David Lo, and Gail C. Murphy. 2021. "A Large Scale Study of Long-Time Contributor Prediction for GitHub Projects." *IEEE Transactions on Software Engineering*, 47(6): 1277–1298.

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts. 2012. "Does Management Matter? Evidence from India *." *The Quarterly Journal of Economics*, 128(1): 1–51.

Bloom, Nicholas, Sadun Rafaella, and John Van Reenen. 2017. "Management as a Technology?" *NBER Working Papers*. Accessed November 2025.

Boin, Arjen, and Michel J. G. van Eeten. 2013. "The Resilient Organization." *Public Management Review*, 15(3): 429–445.

Bonakdarpour, Mohsen, Karen Campbell, Davide Bonaglia, Claire Rousseau, and Sacha Wunsch-Vincent. 2021. "Global Software Spending Surges to Close to USD 700 Billion in 2024, up 50% From 2020; the United States Extends its Lead." Accessed November 2025.

Bosu, Amiangshu, and Jeffrey C. Carver. 2014. "Impact of Developer Reputation on Code Review Outcomes in OSS projects: An Empirical Investigation." *ESEM '14*, 1–10. New York, NY, USA.

Bruhn, Miriam, Dean Karlan, and Antoinette Schoar. 2018. "The Impact of Consulting Services on Small and Medium Enterprises: Evidence from a Randomized Trial in Mexico." *Journal of Political Economy*, 126(2): 635–687.

Carmeli, Abraham, and Gideon D. Markman. 2011. "Capture, Governance, and Resilience: Strategy Implications from the History of Rome." *Strategic Management Journal*, 32(3): 322–341.

Constantinou, Eleni, and Tom Mens. 2017. "An Empirical Comparison of Developer Retention in the RubyGems and npm Software Ecosystems." *Innovations in Systems and Software Engineering*, 13(2): 101–115.

Conti, Annamaria, Christian Peukert, and Maria Roche. 2025. "Beefing IT Up for Your Investor? Engagement with Open Source Communities, Innovation, and Startup Funding: Evidence from GitHub." *Organization Science*, 36(4): 1551–1573.

Conti, Annamaria, Vansh Gupta, Jorge Guzman, and Maria P. Roche. 2023. "Incentivizing Innovation in Open Source: Evidence from the GitHub Sponsors Program."

Crichton, Margaret T., Cameron G. Ramsay, and Terence Kelly. 2009. "Enhancing Organizational Resilience Through Emergency Planning: Learnings from Cross-Sectoral Lessons." *Journal of Contingencies and Crisis Management*, 17(1): 24–37.

Cross, Rob. 2020. "When Collaboration Fails and How to Fix It." *MIT Sloan Management Review*.

Crowston, Kevin, and James Howison. 2006. "Hierarchy and Centralization in Free and Open Source Software Team Communications." *Knowledge, Technology & Policy*, 18(4): 65–85.

Danneels, Erwin. 2008. "Organizational Antecedents of Second-Order Competences." *Strategic Management Journal*, 29(5): 519–543.

DesJardine, Mark, Pratima Bansal, and Yang Yang. 2019. "Bouncing Back: Building Resilience Through Social and Environmental Practices in the Context of the 2008 Global Financial Crisis." *Journal of Management*, 45(4): 1434–1460. Publisher: SAGE Publications Inc.

Dietrichson, Jens, Jens Gudmundsson, and Torsten Jochem. 2022. "Why don't we talk about it? Communication and coordination in teams." *Journal of Economic Behavior & Organization*, 197: 257–278.

Duchek, Stephanie. 2020. "Organizational Resilience: A Capability-Based Conceptualization." *Business Research*, 13(1): 215–246.

Dutton, Jane. 2010. "The Heart of Reuters (A)." Accessed August 2025.

Fiol, C. Marlene, and Marjorie A. Lyles. 1985. "Organizational Learning." *The Academy of Management Review*, 10(4): 803–813.

Freyaldenhoven, Simon, Christian Hansen, Jorge Pérez Pérez, and Jesse M. Shapiro. 2025. "Visualization, Identification, and Estimation in the Linear Panel Event-Study Design." Accessed August 2025.

Github. 2025*a*. "Adding a Code of Conduct to your Project." Accessed October 2025.

Github. 2025*b*. "Github GraphQL API." Accessed August 2025.

Github Archive. 2025. "Github Archive." Accessed August 2025.

Gittell, Jody Hoffer, Kim Cameron, Sandy Lim, and Victor Rivas. 2006. "Relationships, Layoffs, and Organizational Resilience: Airline Industry Responses to September 11." *The Journal of Applied Behavioral Science*, 42(3): 300–329. Publisher: SAGE Publications Inc.

Gortmaker, Jeff. 2025. "Open source software policy in industry equilibrium."

Gressgård, Leif Jarle, and Kåre Hansen. 2015. "Knowledge exchange and learning from failures in distributed environments: The role of contractor relationship management and work characteristics." *Reliability Engineering & System Safety*, 133: 167–175.

Hannon, J., and S. Westlund. 2008. "Retaining Talent: Assessing Relationships Among Project Leadership styles, Software Developer Job Satisfaction, and Turnover Intentions." *Journal of Information Technology Management*, XIX(4).

Hillmann, Julia, and Edeltraud Guenther. 2021. "Organizational Resilience: A Valuable Construct for Management Research?" *International Journal of Management Reviews*, 23(1): 7–44.

Hoffer Gittell, Jody. 2002. "Coordinating Mechanisms in Care Provider Groups: Relational Coordination as a Mediator and Input Uncertainty as a Moderator of Performance Effects." *Management Science*, 48(11): 1408–1426.

Hoffmann, Manuel, Frank Nagle, and Yanuo Zhou. 2024. "The Value of Open Source Software." Accessed January 2024.

Home III, John F. 1997. "The Coming Age of Organizational Resilience." *Business Forum*, 22(2/3/4): 24. Publisher: California State University, Los Angeles.

Home III, John F., and John E. Orr. 1997. "Assessing behaviors that create resilient organizations." *Employment Relations Today*, 24(4): 29–39.

Hutto, C., and Eric Gilbert. 2014. "VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text." *Proceedings of the International AAAI Conference on Web and Social Media*, 8(1): 216–225.

Iaffaldano, Giuseppe, Igor Steinmacher, Fabio Calefato, Marco Gerosa, and Filippo Lanubile. 2019. "Why do Developers Take Breaks from Contributing to OSS Projects? A Preliminary Analysis." *SoHeal '19*, 9–16. IEEE Press.

Ichniowski, Casey, Kathryn Shaw, and Giovanna Prennushi. 1997. "The Effects of Human Resource Management Practices on Productivity: A Study of Steel Finishing Lines." *The American Economic Review*, 87(3): 291–313.

Izquierdo-Cortazar, D., G. Robles, F. Ortega, and J.M. Gonzalez-Barahona. 2009. "Using Software Archaeology to Measure Knowledge Loss in Software Projects Due to Developer Turnover." 1–10. ISSN: 1530-1605.

Jansen, Paul. 2025. "TIOBE Index for July 2025." Accessed July 2025.

Jaravel, Xavier, Neviana Petkova, and Alex Bell. 2018. "Team-Specific Capital and Innovation." *American Economic Review*, 108(4-5): 1034–1073.

Jäger, Simon, Jörg Heining, and Nathan Lazarus. 2024. "How Substitutable Are Workers? Evidence from Worker Deaths."

Kantur, Deniz, and Arzu İşeri Say. 2012. "Organizational Resilience: A Conceptual Integrative Framework." *Journal of Management & Organization*, 18(6): 762–773.

Kendra, James M., and Tricia Wachtendorf. 2003. "Elements of Resilience After the World Trade Center Disaster: Reconstituting New York City's Emergency Operations Centre." *Disasters*, 27(1): 37–53.

Khanna, Rajat. 2021. "Aftermath of a Tragedy: A Star's Death and Coauthors' Subsequent Productivity." *Research Policy*, 50(2): 104159.

Kristensen, Ida. 2025. "Building a Resilient, Future-Ready Organization." Accessed August 2025.

Lamorgese, Andrea, Megha Patnaik, Andrea Linarello, and Fabiano Schivardi. 2024. "Management Practices and Resilience to Shocks: Evidence from COVID-19." *Management Science*, 70(12): 8217–9119.

Lawson, Adrienn, Stephen Hendrick, Nancy Rausch, Jeffrey Sica, and Marco Gerosa. 2024. "Shaping the Future of Generative AI."

Lengnick-Hall, Cynthia A., Tammy E. Beck, and Mark L. Lengnick-Hall. 2011. "Developing a Capacity for Organizational Resilience Through Strategic Human Resource Management." *Human Resource Management Review*, 21(3): 243–255.

Lerner, Josh, and Jean Tirole. 2001. "The open source movement: Key research questions." *European Economic Review*, 45(4-6): 819–826. Publisher: Elsevier.

Lerner, Josh, and Jean Tirole. 2002. "Some Simple Economics of Open Source." *The Journal of Industrial Economics*, 50(2): 197–234. Publisher: Wiley.

Liao, Christopher. 2025. "Generalized Random Forest Event Studies." Accessed November 2025.

Luthans, Fred, James B. Avey, Bruce J. Avolio, Steven M. Norman, and Gwendolyn M. Combs. 2006. "Psychological Capital Development: Toward a Micro-Intervention." *Journal of Organizational Behavior*, 27(3): 387–393.

Makkonen, Hannu, Mikko Pohjola, Rami Olkkonen, and Aki Koponen. 2014. "Dynamic Capabilities and Firm Performance in a Financial Crisis." *Journal of Business Research*, 67(1): 2707–2719.

Mallak, Larry A. 1998. "Measuring Resilience in Health Care Provider Organizations." *Health Manpower Management*, 24(4): 148–152.

Maor, Dana, and Michael Park. 2023. "Foster Organizational Resilience for a Stronger 2023." Accessed August 2025.

Maor, Dana, Michael Park, and Brooke Weddle. 2022. "Raising the Resilience of your Organization." Accessed August 2025.

McManus, Sonia, Erica Seville, John Vargo, and David Brunsdon. 2008. "Facilitated Process for Improving Organizational Resilience." *Natural Hazards Review*, 9(2): 81–90.

Milgrom, Paul, and John Roberts. 1990. "The Economics of Modern Manufacturing: Technology, Strategy, and Organization." *The American Economic Review*, 80(3): 511–528.

Milgrom, Paul, and John Roberts. 1995. "Complementarities and fit strategy, structure, and organizational change in manufacturing." *Journal of Accounting and Economics*, 19(2): 179–208. Organizations, Incentives, and Innovation.

Miller, Courtney, David Gray Widder, Christian Kästner, and Bogdan Vasilescu. 2019. "Why Do People Give Up FLOSSing? A Study of Contributor Disengagement in Open Source." 116–129. Cham:Springer International Publishing.

Mockus, Audris, Roy T. Fielding, and James D. Herbsleb. 2002. "Two Case Studies of Open Source Software Development: Apache and Mozilla." *ACM Trans. Softw. Eng. Methodol.*, 11(3): 309–346.

Nagle, Frank. 2017. "Open Source Software and Firm Productivity."

Nassif, Mathieu, and Martin P. Robillard. 2017. "Revisiting Turnover-Induced Knowledge Loss in Software Projects." 261–272.

Oettl, Alexander. 2012. "Reconceptualizing Stars: Scientist Helpfulness and Peer Performance." *Management Science*, 58(6): 1122–1140. Publisher: INFORMS.

Ortiz-de Mandojana, Natalia, and Pratima Bansal. 2016. "The Long-term Benefits of Organizational Resilience through Sustainable Business Practices." *Strategic Management Journal*, 37(8): 1615–1631.

Pal, Rudrajeet, Håkan Torstensson, and Heikki Mattila. 2014. "Antecedents of Organizational Resilience in Economic Crises—an Empirical Study of Swedish Textile and Clothing SMEs." *International Journal of Production Economics*, 147: 410–428. Building Supply Chain System Capabilities in the Age of Global Complexity: Emerging Theories and Practices.

Rambachan, Ashesh, and Jonathan Roth. 2023. "A More Credible Approach to Parallel Trends." *The Review of Economic Studies*, 90(5): 2555–2591.

Rashid, Mehvish, Paul M. Clarke, and Rory V. O'Connor. 2017. "Exploring Knowledge Loss in Open Source Source (OSS) Projects." 481–495. Cham:Springer International Publishing.

Rashid, Mehvish, Paul M. Clarke, and Rory V. O'Connor. 2019. "A Systematic Examination of Knowledge Loss in Open Source Software Projects." *International Journal of Information Management*, 46: 104–123.

Rigby, Peter C., Yue Cai Zhu, Samuel M. Donadelli, and Audris Mockus. 2016. "Quantifying and Mitigating Turnover-induced Knowledge Loss: Case Studies of Chrome and a Project at Avaya." *ICSE '16*, 1006–1016. New York, NY, USA:Association for Computing Machinery.

Robert, Benoît. 2009. *Organizational Resilience: Concepts and Evaluation Method.* Presses internationales Polytechnique.

Robles, Gregorio, Jesus M Gonzalez-Barahona, and Martin Michlmayr. 2005. "Evolution of Volunteer Participation in Libre Software Projects: Evidence from Debian." *Proceedings of the First International Conference on Open Source Systems*, 100–107.

Sahebjamnia, N., S. A. Torabi, and S. A. Mansouri. 2015. "Integrated Business Continuity and Disaster Recovery Planning: Towards Organizational Resilience." *European Journal of Operational Research*, 242(1): 261–273.

Sheffi, Yossi, and James B. Rice Jr. 2005. "A Supply Chain View of the Resilient Enterprise." *MIT Sloan Management Review.* Accessed August 2025.

Steinmacher, Igor, Christoph Treude, and Marco Aurelio Gerosa. 2019. "Let Me In: Guidelines for the Successful Onboarding of Newcomers to Open Source Projects." *IEEE Software*, 36(4): 41–49. Conference Name: IEEE Software.

Suarez, Fernando F., and Juan S. Montes. 2020. "Building Organizational Resilience." *Harvard Business Review.* Accessed August 2025.

Sun, Liyang, and Sarah Abraham. 2021. "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects." *Journal of Econometrics*, 225(2): 175–199. Themed Issue: Treatment Effect 1.

Tang, David. 2025. "Business Resilience Initiative for Cloud Hosting Service in Tech Market." Accessed August 2025.

Taylor, Frederick Winslow. 1911. *The Principles of Scientific Management.* NuVision Publications, LLC.

The Python Software Foundation. 2025. "PyPI Download Stats." Accessed October 2025.

Uysal, S. Derya. 2015. "Doubly Robust Estimation of Causal Effects with Multivalued Treatments: An Application to the Returns to Schooling." *Journal of Applied Econometrics*, 30(5): 763–786.

von Hippel, Eric, and Georg von Krogh. 2003. "Open Source Software and the "Private-Collective" Innovation Model: Issues for Organization Science." *Organization Science*, 14(2): 209–223. Publisher: INFORMS.

von Krogh, Georg, Sebastian Spaeth, and Karim R Lakhani. 2003. "Community, Koining, and Specialization in Open Source Software Innovation: A Case Study." *Research Policy*, 32(7): 1217–1241.

W3Cook. 2015. "OS Usage Trends and Market Share." Accessed August 2025.

Waldinger, Fabian. 2010. "Quality Matters: The Expulsion of Professors and the Consequences for PhD Student Outcomes in Nazi Germany." *Journal of Political Economy*, 118(4): 787–831.

Wikipedia. 2025. "Great Resignation." *Wikipedia, The Free Encyclopedia.* Accessed November 2025.

Woodward, Joan. 1965. *Management and Technology. Problems of progress in industry*, H.M. Stationery Office.

Xu, Bo, and Donald R. Jones. 2010. "Volunteers' Participation in Open Source Software Development: A Study from the Social-relational Perspective." *SIGMIS Database*, 41(3): 69–84.

Yu, Yiqing, Alexander Benlian, and Thomas Hess. 2012. "An Empirical Study of Volunteer Members' Perceived Turnover in Open Source Software Projects." 3396–3405. ISSN: 1530-1605.

A   Additional Empirical Analysis

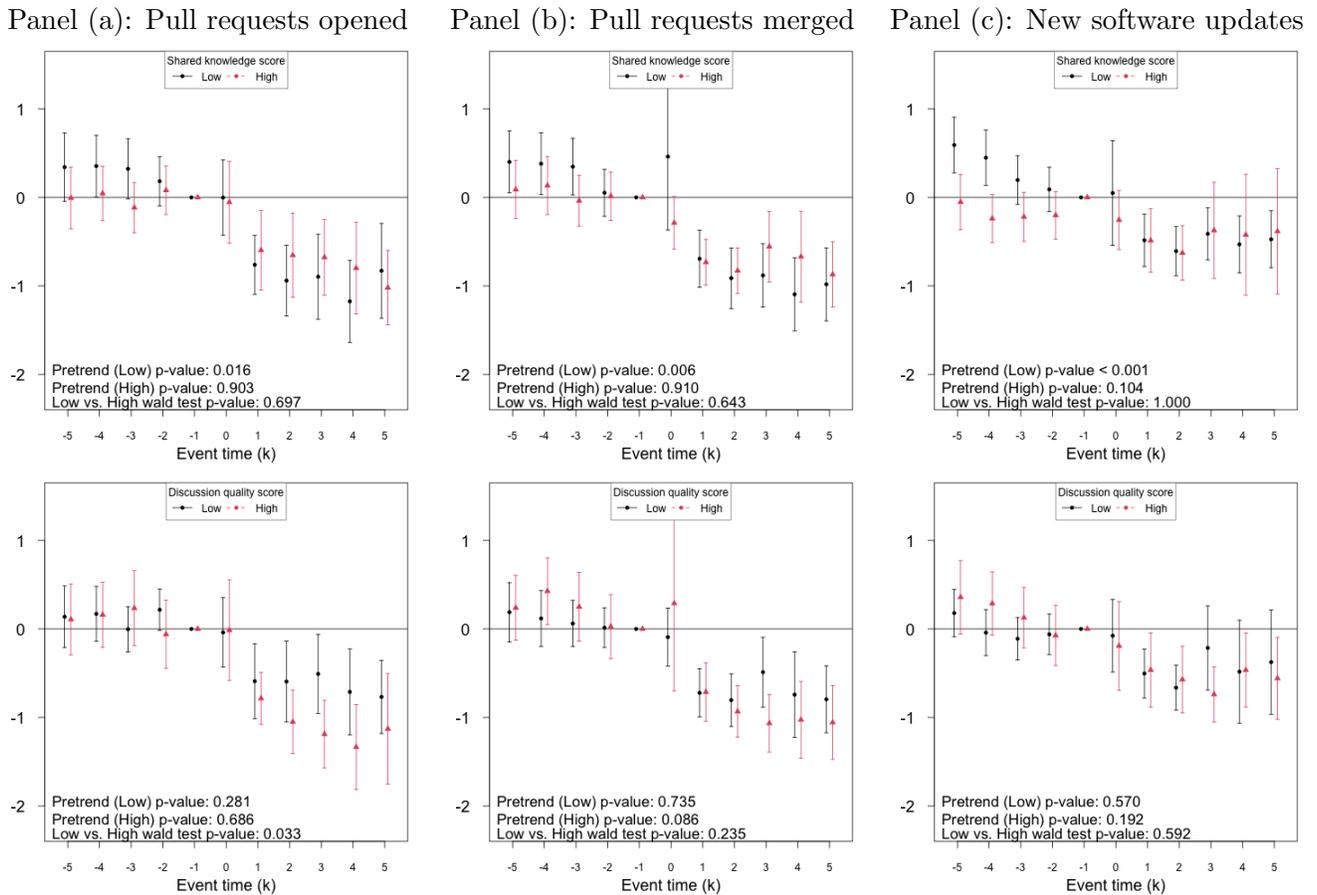Figure A1: Organizational Resilience by Organizational Practice Category Scores

Panel (a): Pull requests opened   Panel (b): Pull requests merged   Panel (c): New software updates

Figure notes: The first row of estimates in Panel (a) presents separate event study estimates of the impact of key member departures on the number of pull requests opened for organizations with high and low knowledge level scores. Organizations are classified as having high knowledge level scores if their knowledge score is above the cross-organization median, and low otherwise. Outcomes are standardized using each organization's mean and standard deviation over the five pre-periods, and each period spans six months. Estimates, expressed in standard-deviation units, are obtained using Sun and Abraham (2021). Confidence intervals are 95% intervals based on asymptotic standard errors clustered at the organization level. The two reported pre-trend p-values are Wald tests of the null that all pre-period coefficients equal zero for the high- and low-collaboration subsets. The final p-value is from a Wald test of the null that post-period event-study coefficients do not differ between the two subsets. Panels (b) and (c) repeat this analysis using the number of pull requests merged and the number of new software releases as outcomes. The second row of estimates replaces the collaboration score with the discussion quality score.

Figure A2: Organizational Resilience by Organizational Practice Category Scores

Panel (a): Pull requests opened   Panel (b): Pull requests merged   Panel (c): New software updates
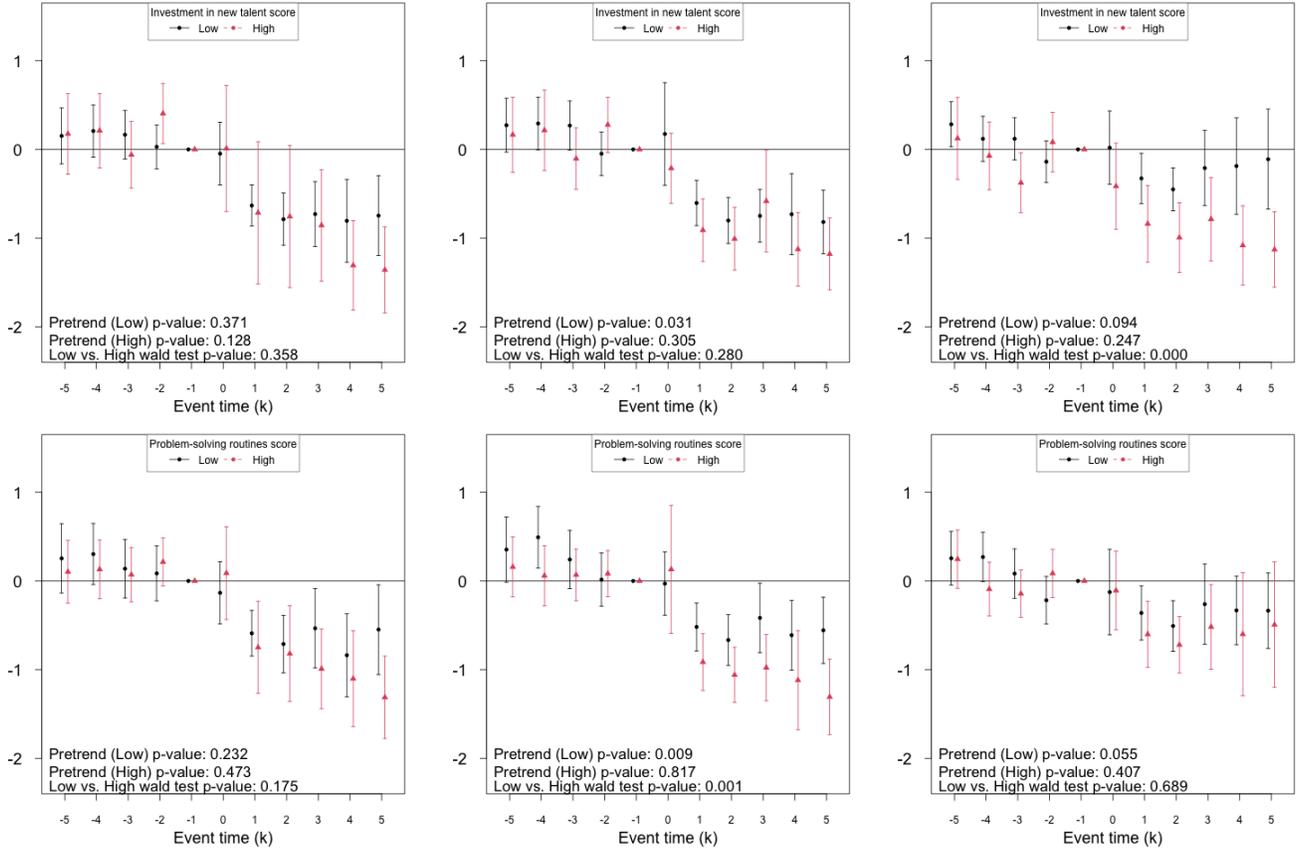

Figure notes: The first row of estimates in Panel (a) presents separate event study estimates of the impact of key member departures on the number of pull requests opened for organizations with high and low knowledge level scores. Organizations are classified as having high knowledge level scores if their knowledge score is above the cross-organization median, and low otherwise. Outcomes are standardized using each organization's mean and standard deviation over the five pre-periods, and each period spans six months. Estimates, expressed in standard-deviation units, are obtained using Sun and Abraham (2021). Confidence intervals are 95% intervals based on asymptotic standard errors clustered at the organization level. The two reported pre-trend p-values are Wald tests of the null that all pre-period coefficients equal zero for the high- and low-collaboration subsets. The final p-value is from a Wald test of the null that post-period event-study coefficients do not differ between the two subsets. Panels (b) and (c) repeat this analysis using the number of pull requests merged and the number of new software releases as outcomes. The second row of estimates replaces the collaboration score with the discussion quality score