Public Signaling and Knowledge Spillovers Before Job Transitions

Jung Ho Choi
Stanford University
Graduate School of Business

Dan Li
Singapore Management University
School of Accountancy

Aluna Wang
HEC Paris

May 2026

Preliminary Draft. Please Don't Distribute.

ABSTRACT

Labor-market frictions can induce jobseekers to publicly disclose their skills, creating knowledge spillovers before those workers actually move. We study this mechanism using linked data on GitHub activity and employment histories for more than 22,000 software engineers and 46,637 job transitions in our event-window sample from 2011 to 2021. We find that engineers increase public GitHub activity in the months preceding a job transition. We find more qualified evidence that this activity is taken up by other developers. The increase is weaker when signaling incentives appear lower (e.g., among senior and higher-compensation engineers and during tight labor markets) and when those public activities might entail revealing sensitive information (e.g., those at R&D-intensive firms). External validation of jobseekers' GitHub activities, especially forks received, is associated with subsequent compensation gains and moves to larger firms, while simple signaling volume is negatively associated with career transitions. Using firm-level headcount contractions as an instrument for worker departure, we find that contraction-induced departures lead to economically meaningful increases in worker-created public coding activity. We further show that firms with greater engineer disclosure are less likely to maintain positive earnings changes persistently, suggesting a proprietary cost of their employees' GitHub activities. These findings reveal a pre-mobility channel of knowledge spillovers: labor market frictions affect not only where workers move, but also what knowledge becomes publicly available before they move.

Keywords: Disclosure, Labor Mobility, Signaling, Knowledge Spillover, Open-source Software

JEL Codes: M41, M14, M51, J71, J78, G38

Jung Ho Choi: jungho@stanford.edu. Dan Li: danli@smu.edu.sg. Aluna Wang: wanga@hec.fr. We thank Ed deHaan and Gurpal Sran for helpful comments and suggestions. This study has been exempted from the IRB review by SMU.

1. Introduction

Knowledge spillovers are central to economic growth and firm performance. A large literature in economics, strategy, and accounting identifies labor mobility as an important mechanism through which knowledge diffuses across firms: workers carry skills, routines, and tacit knowledge from one employer to another, shaping innovation, productivity, and organizational outcomes (Arrow 1962; Jaffe, Trajtenberg, and Henderson 1993; Almeida and Kogut 1999; Rosenkopf and Almeida 2003; Song, Almeida, and Wu 2003; Tambre and Hitt, 2013; Barrios and Gallemore 2024). This view is especially important in knowledge-intensive industries, where human capital is difficult to codify and where firms rely heavily on highly skilled employees. Yet the standard view of labor-market knowledge spillovers typically begins at the moment of mobility: workers leave one firm, join another, and transfer knowledge through their new employment relationship. In this paper, we examine whether knowledge spillovers begin earlier, before workers actually move.

Our central argument is that labor market frictions can generate anticipatory knowledge spillovers by inducing workers to publicly demonstrate their skills and capabilities. When employers have imperfect information about worker ability, job-seeking workers have incentives to produce credible signals of quality (Spence 1973; Greenwald 1986). In many labor markets, these signals are private or only partially observable, such as referrals, interviews, or résumés (Burks et al. 2015; Hung et al. 2023). In software engineering, however, workers can signal through public digital artifacts, including code repositories, commits, and pull requests. These artifacts credibly reveal technical ability to prospective employers, and much of that credibility comes from outside the worker. Because other developers fork, reuse, and build on code only when it is genuinely useful, third-party adoption acts as a crowdsourced, hard-to-fake verification of

quality: a signal the worker cannot manufacture alone, precisely because it depends on others choosing to use it. But this is exactly what makes the signal double-edged: the adoption that validates it is the transfer of knowledge. The same publicly observable activity that helps workers overcome labor-market frictions discloses knowledge to competitors, collaborators, and the broader developer community.

This mechanism highlights a tension that is central to disclosure theories. Public disclosure can improve matching by reducing information asymmetry, but it can also impose disclosure costs by revealing valuable knowledge to outside parties (Verrecchia 1983; Dye 1985; Darrough and Stoughton 1990; Kim and Valentine 2021). For workers, public coding activity may increase labor-market visibility and improve outside opportunities. For firms, however, such activity may expose technical knowledge, reveal the capabilities of their employees, or facilitate employee poaching. The resulting disclosure decision is therefore not merely a labor-market choice; it is also a knowledge-transfer decision that reaches beyond the worker and the prospective employer. By publicly demonstrating skill, a worker confers benefits on third parties who reuse the disclosed artifacts, while potentially imposing costs on the current employer whose know-how becomes visible to competitors. This setting lets us study how individually optimal signaling generates externalities and to measure both sides directly: the downstream reuse of disclosed code, and the consequences for the origin firm.

Software engineers provide a powerful setting for studying this mechanism. Their output is often digital, modular, and publicly observable when produced on open-source platforms. Prior work argues that open-source contributions can serve as credible signals of programming ability and career concerns because individual contributions are visible, verifiable, and difficult to fake (Lerner and Tirole 2002; Marlow and Dabbish 2013; Abou El-Komboz and Goldbeck 2025). This

literature, however, largely infers the career-concerns motive from the visibility of contributions rather than observing whether such signaling is actually tied to specific job moves: a test that requires linking individual contribution histories to employer-to-employer transitions, which is the central data advantage we bring. GitHub, in particular, allows developers to demonstrate technical skill through timestamped repositories, commits, pull requests, and external engagement. At the same time, GitHub activity can generate real knowledge diffusion: other developers can fork repositories, build on existing code, submit improvements, and incorporate disclosed solutions into their own work. Prior research on open-source software shows that collaboration networks and public code repositories facilitate both direct and indirect knowledge spillovers (Fershtman and Gandal 2011). This setting therefore allows us to observe both sides of the mechanism: workers' public skill demonstration and outsiders' use of the disclosed knowledge.

We examine this mechanism using linked data on GitHub activity and employment histories. Our main event-window sample consists of 22,293 software engineers and 46,637 employer-to-employer transitions from 2011 to 2021. We also use a larger position-level transition cross-section of 54,037 transitions for match-quality analyses that do not require construction of the engineer-day event window. For each engineer in the event-window sample, we observe daily public GitHub activity around job transitions, including commits, pull requests, repositories created, forks received, active forks, and pull requests merged. We link these activity measures to employment histories, compensation estimates, seniority, and employer characteristics, allowing us to examine how public coding activity changes before and after job transitions and whether these changes are associated with labor-market and firm-level outcomes. This design builds on recent work using GitHub to study knowledge work, open-source production, and labor-market behavior (McDermott and Hansen 2021; Abou El-Komboz and Goldbeck 2025). Unlike these

studies, which observe GitHub activity in isolation, we link it to an employer-employee matched dataset, letting us connect public contributions to actual job transitions.

We first document that engineers increase public GitHub activities in the months preceding a job transition. The increase is concentrated in activities that are visible to outside observers and that can plausibly serve as signals of technical ability, such as commits and repository creation. This pre-move increase is followed by a sharp decline after the transition, suggesting that the activity is not simply a persistent increase in productivity or platform use. Instead, the timing is consistent with workers increasing public output when the labor-market value of signaling is highest. This interpretation is consistent with models of career concerns and signaling, in which agents exert observable effort when reputational benefits are greatest (Holmström 1999; Chevalier and Ellison 1999), and with recent evidence that open-source activity increases during job-search periods (Abou El-Komboz and Goldbeck 2025).

We next examine whether those signaling activities generate knowledge diffusion. Public disclosure does not necessarily imply spillovers; disclosed knowledge must be noticed, used, or extended by others. We therefore examine external engagement with engineers' work using measures such as forks, active forks, and merged pull requests. These outcomes capture whether other developers use or build upon the engineer's publicly disclosed code. The evidence indicates that this pre-move activity is not merely cosmetic profile-padding: external use of disclosed work is higher before job transitions than after, consistent with the idea that labor-market signaling can generate knowledge spillovers before mobility occurs. This distinction between disclosure and actual utilization is important because knowledge spillovers require not only public availability but also transmission, recombination, or use by other agents (Jaffe, Trajtenberg, and Henderson 1993; Fershtman and Gandal 2011; Barrios and Gallemore 2024).

We then examine whether these GitHub activities vary with signaling incentives and the sensitivity of the underlying knowledge. The signaling interpretation predicts weaker responses when workers already possess strong labor-market signals and stronger responses when workers have greater need to demonstrate ability. Consistent with this prediction, pre-move activity is weaker among senior engineers and higher-compensation engineers, and during tight labor markets, when workers may need less additional signaling to attract employers. A second prediction concerns what is disclosed: when the underlying knowledge is more sensitive, workers should be more reluctant to expose it publicly, even at the cost of a weaker signal. Consistent with this, pre-move activity is weaker among engineers at R&D-intensive firms, where the knowledge involved is more proprietary and workers are correspondingly discouraged from disclosing it. These cross-sectional patterns help distinguish the signaling mechanism from alternative explanations such as general career progression, changes in workload, or mechanical changes in GitHub use. They also connect the paper to the disclosure literature's central trade-off between information benefits and the cost of revealing proprietary knowledge, as well as to labor-market signaling models in which incentives vary with reputation and outside options (Spence 1973; Verrecchia 1983; Greenwald 1986; Kim and Valentine 2021).

We also study whether GitHub activities during job search periods are associated with subsequent labor-market outcomes, but we interpret these tests more carefully than the signaling event studies. If public GitHub activity serves as a credible signal, externally validated activity should be valued more strongly than raw activity volume, because validation reflects not only the worker's decision to disclose but also the market's response to it. The evidence supports this distinction. Raw signaling volume is, if anything, negatively associated with favorable career transitions: engineers who simply produce more public activity before moving tend to see smaller

compensation gains and are less likely to move to larger firms. By contrast, externally validated activity, especially forks received, is positively associated with both compensation gains and moves to larger firms. The labor market thus appears to reward public disclosure when outside users validate the disclosed work, rather than treating all visible effort as equally informative; unvalidated volume may even signal weaker quality. This evidence relates to studies showing that public performance records, online reputation, and verifiable signals can improve labor-market outcomes by reducing information frictions (Pallais 2014; Agrawal, Lacetera, and Lyons 2016; Gow, Wahid, and Yu 2018).

A central empirical concern is selection: workers who increase GitHub activity before moving may differ systematically from those who do not. More ambitious or more able engineers, for example, may both signal more and obtain better jobs. Such selection cannot fully explain the event-time dynamics we document, a sharp pre-move rise followed by a decline after the move, since fixed differences in ability or ambition would not produce activity that spikes precisely around the transition. It could nonetheless confound the cross-sectional and career-outcome results. To address this concern, we use firm-level headcount contractions as an instrument for worker departure. The identifying idea is that firm-level contractions raise the probability of departure for reasons largely unrelated to an individual engineer's latent signaling propensity. Using this design, we find that contraction-induced departures lead to economically meaningful increases in those workers' public software projects, and that the response strengthens with the severity of the contraction. These results support the interpretation that employer-side shocks increase mobility incentives, which in turn induce public skill demonstration.

Finally, we examine whether worker disclosure has implications for firms. When engineers publicly disclose technical output, that knowledge can be redistributed outside the firm, learned

from, reused, or built upon by outside parties, eroding the persistence of the firm's competitive advantage. Consistent with this proprietary-cost channel, we find that firms with greater engineer disclosure are less likely to sustain positive earnings changes persistently. These firm-level analyses are associational and should be interpreted as evidence on broader consequences of disclosure rather than as definitive causal estimates. This analysis connects the paper to research on human capital as a source of firm performance and to studies showing that labor mobility transfers economically relevant knowledge across firms (Becker 1964; Zander and Kogut 1995; Rosenkopf and Almeida 2003; Song, Almeida, and Wu 2003; Barrios and Gallemore 2024). Barrios and Gallemore (2024), for example, emphasize that labor markets can transfer specialized knowledge across firms and that the value of such knowledge depends on the fit between the origin and destination environments.

Our paper makes three contributions. First, we contribute to the literature on labor mobility and knowledge diffusion by showing that knowledge spillovers can arise before workers move. Prior work emphasizes the transfer of knowledge after employees join new firms; we show that the anticipation of mobility can itself induce knowledge disclosure. Second, we contribute to the literature on labor-market signaling by showing that signaling can generate real externalities. In our setting, public signals are not merely informative; they are productive artifacts that others can use. Third, we contribute to the accounting and disclosure literature by identifying a worker-level disclosure channel that affects both labor-market matching and firm-level outcomes. The findings suggest that labor-market frictions shape not only who moves across firms, but also what knowledge becomes publicly available before the move occurs.

2. Theoretical Foundation

2.1 Labor Mobility, Information Frictions, and Skill Signaling

Labor mobility is an important mechanism through which knowledge is reallocated across firms. When workers move, they carry skills, routines, and experience from one organizational environment to another, facilitating the diffusion of productive knowledge across firm boundaries (Arrow 1962; Almeida and Kogut 1999; Rosenkopf and Almeida 2003; Song, Almeida, and Wu 2003; Barrios and Gallemore 2024). This channel is especially important in knowledge-intensive industries, where productive capabilities are embedded in individuals and are difficult for firms to fully codify, contract over, or retain.

At the same time, labor mobility occurs in markets characterized by substantial information frictions. Prospective employers do not perfectly observe worker ability, and workers often have limited direct access to the full set of potential employers. These frictions create incentives for workers to engage in costly signaling to credibly communicate their skills (Spence 1973; Greenwald 1986). In traditional labor markets, such signals include education, prior employers, referrals, interviews, and professional certifications. These signals are informative, but they are often coarse. They reveal broad credentials rather than direct evidence of task-specific ability.

Digital production platforms change the nature of this signaling problem (Lerner and Tirole, 2002). Workers can now make aspects of their productive output publicly observable. For software engineers, public code repositories, commits, pull requests, and related artifacts provide verifiable evidence of technical ability. These signals are difficult to fake because they require actual production effort and can be evaluated by other developers. They also allow workers to communicate capabilities to a broader set of potential employers than would be possible through private search alone.

The incentive to produce public signals should be particularly strong around job transitions. As workers approach the labor market, the marginal benefit of making their output visible

increases because public evidence of skill may attract employer attention, improve bargaining power, and facilitate better matching. In this sense, public coding activity can serve as a decentralized labor-market signal: it allows workers to demonstrate ability without relying entirely on résumés, referrals, or current-employer reputation.

2.2 Public Skill Signals as Disclosure

Although public coding activity can serve as a labor-market signal, it also has the economic properties of disclosure. The worker reveals information that was previously private or less visible: technical ability, problem-solving style, project experience, and, in some cases, knowledge embedded in code. This disclosure can reduce information asymmetry in the labor market, but it can also reveal knowledge to parties beyond the intended audience.

This trade-off parallels a central theme in the disclosure literature. Disclosure can improve market efficiency by reducing information asymmetry, but it can also impose proprietary costs when disclosed information benefits competitors or weakens private advantage (Verrecchia 1983; Dye 1985; Darrough and Stoughton 1990; Kim and Valentine 2021). In our setting, the disclosing agent is the worker rather than the firm. The private benefit accrues primarily to the worker through improved labor-market visibility, while the potential cost may be borne partly by the current employer if public output reveals knowledge, methods, or capabilities that were previously less accessible.

This distinction is important because worker disclosure need not be aligned with firm preferences. A worker preparing for the labor market may have incentives to make technical output visible, even when the current employer would prefer to keep certain knowledge less observable. Thus, public skill signaling creates a wedge between worker-level labor-market incentives and firm-level proprietary-cost concerns. The disclosure decision is not merely about whether the

worker produces knowledge, but about whether that knowledge becomes observable to outside parties.

The software engineering setting makes this tension empirically observable. GitHub activity is public, timestamped, and visible to prospective employers, but it is also visible to other developers, firms, and competitors. A repository can demonstrate the engineer's ability while simultaneously making code, design choices, and technical solutions available to the broader market. The same activity can therefore improve labor-market matching and redistribute knowledge outside the boundaries of the current firm.

2.3 Pre-Mobility Knowledge Spillovers

The standard view of labor-market knowledge spillovers emphasizes what happens after mobility: workers transfer knowledge to new employers once they move. Our framework highlights a distinct pre-mobility channel. When workers publicly disclose output in order to signal ability, knowledge can spill over before the worker changes firms.

Public availability alone, however, is not sufficient for knowledge diffusion. Disclosed knowledge must be noticed, used, extended, or incorporated by others. In the GitHub setting, such utilization can be observed through external engagement. Other developers can fork repositories, build on disclosed code, submit pull requests, or merge contributions into existing projects. These actions indicate that the disclosed output is not merely visible but is being used by other market participants.

This mechanism differs from traditional mobility-based diffusion in two ways. First, the timing differs: knowledge becomes available before the worker joins a new employer. Second, the audience differs: the knowledge is not transferred only to the destination firm, but potentially to a

broad set of outside developers, firms, and competitors. The spillover is therefore a byproduct of a private signaling decision rather than the direct result of employment at the destination firm.

Prior research on open-source software shows that public code and collaboration networks can facilitate direct and indirect knowledge spillovers (Fershtman and Gandal 2011). Our setting extends this idea by linking such spillovers to labor-market incentives. If workers increase public output before job transitions, and if other developers engage with that output, then labor-market frictions can generate knowledge spillovers even before physical worker mobility occurs.

This pre-mobility channel also clarifies why signaling activity can have broader economic consequences. In many signaling models, the signal is socially wasteful unless it improves matching (Spence 1973). In this setting, however, the signal can have productive content. Public code is not only a costly indicator of ability; it is also an artifact that others can inspect, reuse, and modify. Signaling can therefore generate real spillovers even when the worker's primary motive is private career advancement.

2.4 Heterogeneity in Signaling Incentives and Proprietary Costs

The signaling interpretation implies systematic heterogeneity. Workers should increase public disclosure most when the marginal benefit of signaling is high and least when the marginal benefit is low or the proprietary cost is high.

Signaling incentives should be weaker for workers who already possess strong labor-market signals. Senior engineers and higher-compensation engineers likely have more established reputations, richer employment histories, and stronger observable credentials. Because these workers already have credible signals of ability, the incremental benefit of additional public

GitHub activity should be lower. By contrast, less senior or lower-compensation engineers may benefit more from demonstrating ability through observable output.

Labor-market conditions should also matter. When demand for software engineers is high, workers may need less additional signaling to attract employer attention. Conversely, when outside opportunities are more limited, workers may have stronger incentives to differentiate themselves through public output. This variation helps distinguish signaling from alternative explanations such as general career progression or changes in workload. If pre-move GitHub activity simply reflected productivity growth around transitions, it would be less likely to vary systematically with the strength of outside labor-market signals.

Proprietary costs can further moderate disclosure. Engineers at R&D-intensive firms may work in environments where technical knowledge is more valuable and where public disclosure is more sensitive. These workers may face stronger explicit or implicit constraints on what can be shared publicly. Even when they have incentives to signal, the cost of revealing knowledge may dampen public activity. Thus, the same mechanism that generates worker-level benefits can impose costs on the current employer, especially when disclosed knowledge is strategically valuable.

Together, these forces imply that pre-move public disclosure should vary with both the benefits of signaling and the costs of revealing knowledge. This cross-sectional variation is useful empirically because it allows us to distinguish the proposed mechanism from more mechanical explanations for GitHub activity around job transitions.

2.5 Labor-Market Consequences for Workers

If public GitHub activity functions as a credible signal, it should be associated with subsequent labor-market outcomes. Prospective employers may use public coding activity to infer ability, motivation, technical specialization, or fit with the hiring firm's needs. Prior evidence suggests that public performance records and verifiable online signals can reduce information frictions and improve labor-market outcomes, particularly when worker quality is otherwise difficult to observe (Pallais 2014; Agrawal, Lacetera, and Lyons 2016; Gow, Wahid, and Yu 2018).

In our setting, not all public activity should be equally informative. Activity that receives external engagement may be especially valuable because it is validated by other users. A repository that is forked or a contribution that is merged provides a stronger signal than activity that is merely posted but ignored. External engagement therefore helps distinguish visible effort from output that others find useful.

These labor-market consequences are important for interpreting the worker's disclosure decision. If pre-move public activity is associated with better outcomes, then disclosure can be understood as an investment in labor-market visibility. At the same time, the relation between disclosure and outcomes need not be uniformly positive across all dimensions of match quality. Public signaling may help workers obtain higher-paying jobs, but it may also facilitate riskier or more exploratory matches. The central implication is that the labor market appears to value public, verifiable evidence of skill, especially when that evidence is externally validated.

2.6 Firm-Level Consequences: Knowledge Redistribution

Worker disclosure can also have consequences for the firms that employ these engineers. When engineers publicly disclose technical output, they may reveal knowledge that would otherwise remain more closely tied to the firm. This disclosure can erode the firm's private advantage if competitors, customers, or other developers learn from the disclosed work. The effect

does not require formal intellectual property leakage. Even generalizable code, reusable technical approaches, or evidence about employee capabilities can reduce the informational advantages that firms derive from their internal human capital.

These advantages are a source of firm-specific rents, and rents that competitors can observe and imitate are harder to sustain. We therefore expect knowledge redistribution to manifest in the persistence of firm performance: if public disclosure weakens the firm-specific knowledge advantages that allow strong performance to persist, firms with greater engineer disclosure should exhibit lower persistence in performance outcomes. The channel is also likely to operate with a lag, because competitors and other market participants need time to absorb, adapt, and deploy disclosed knowledge before it affects the disclosing firm's results.

3. Data and Sample

3.1 Data Sources

Our analysis combines two primary data sources. The first is GitHub Archive, which records public activity on GitHub, the largest online platform for software development and open-source collaboration. GitHub Archive provides time-stamped event-level data on public repositories, including commits, pull requests, repository creation, forks, and related activities. These data allow us to observe when software engineers produce publicly visible output and when other developers engage with that output.

The second data source is Revelio Labs, an employer–employee matched panel constructed from online professional profiles. Revelio provides detailed employment histories, including employer names, job titles, start and end dates, seniority, and compensation estimates. These data allow us to identify employer-to-employer transitions and to characterize both origin and

destination jobs. We also use Revelio to construct worker-level and employer-level characteristics used in our cross-sectional and outcome analyses.

We link GitHub users to Revelio employment histories at the individual level. The linked data allow us to observe both the timing of job transitions and the evolution of public coding activity around those transitions. This structure is central to our empirical design because it allows us to distinguish public activity that occurs before a worker moves from activity that occurs after the worker joins a new employer.

3.2 Sample Construction

We begin with software engineers who have valid GitHub accounts and employment histories in Revelio (Online Appendix Table A2). We identify employer-to-employer job transitions using changes in the worker's employer recorded in Revelio. For each transition, we define the move date as the start date of the new job. We retain transitions for which we can observe GitHub activity in an event window around the move date.

Our main event-window sample consists of 22,293 unique software engineers and 46,637 employer-to-employer job transitions from January 2011 through March 2021. This sample is constructed after linking Revelio transitions to the GitHub engineer-day event panel. The average engineer in the event-window sample experiences 2.09 observed job transitions. We construct an engineer-by-day event panel centered on each transition and observe activity within a window of up to 360 days before and 360 days after the move date; transitions near the beginning or end of the GitHub sample can have shorter observed windows. This yields 24,936,731 engineer-day observations.

The event-centered structure allows us to compare each engineer's public activity in the months leading up to a job transition with activity farther from the transition and after the move. Because engineers may experience multiple job transitions, our empirical design treats each transition as a separate event and includes event fixed effects in the main specifications. This approach absorbs stable differences across engineer-transition events and focuses the analysis on within-event changes in activity over time.

3.3 Measures of Public Signaling Activity

We construct three measures of public signaling activity. The first is the number of commits, which captures changes or additions to code. Commits are a broad measure of coding activity and provide observable evidence of technical effort. The second is the number of pull requests opened, which captures attempts to contribute code to repositories and is often associated with collaboration or participation in broader projects. The third is the number of repositories created, which captures the initiation of new public projects and represents a more visible form of skill demonstration.

These measures capture different dimensions of public output. Commits reflect the intensity of coding effort. Pull requests opened reflect attempts to contribute to projects that may involve external review. Repositories created reflect deliberate public presentation of a project or technical artifact. Because these activities are visible on GitHub, they can help prospective employers evaluate technical ability, specialization, and engagement with the developer community.

The distributions of these variables are highly skewed, with a large share of engineer-days containing no activity. We therefore use the inverse hyperbolic sine transformation of daily activity

counts in the empirical analyses. This transformation accommodates zero values while preserving variation in the upper tail of the distribution.

3.4 Measures of Knowledge Diffusion

We measure knowledge diffusion using external engagement with the engineer's public GitHub activity. Our primary measures are forks received, active forks, and pull requests merged into the engineer's repositories. These variables capture whether other developers use, extend, or contribute to the engineer's publicly disclosed work.

Forks received capture instances in which another user copies a repository, often to inspect, modify, or build on the code. Active forks refine this measure by focusing on forks that show subsequent activity, providing stronger evidence that the disclosed code is not merely observed but used. Pull requests merged capture contributions from other developers that are accepted into the engineer's repositories, indicating collaboration and integration of outside contributions.

These measures are important because public disclosure alone does not necessarily imply knowledge diffusion. A worker may publish code that no one uses. By examining external engagement, we can distinguish public skill demonstration from activity that is taken up by others. As with the signaling measures, we use inverse hyperbolic sine transformations of daily diffusion outcomes in the empirical analyses.

3.5 Worker, Employer, and Career Outcome Variables

We use Revelio to construct worker and employer characteristics that help explain heterogeneity in signaling incentives and disclosure costs. Worker characteristics include seniority and compensation. Seniority captures the worker's career stage, while compensation provides a proxy for labor-market validation and outside-option strength. Employer characteristics include

firm size, R&D intensity, technology-industry classification, and other financial variables for firms that can be linked to Compustat.

We use these variables to examine whether pre-move public activity varies with the benefits and costs of disclosure. For example, senior and higher-compensation engineers may have less need to signal because they already possess stronger labor-market credentials. Engineers at R&D-intensive firms may face higher disclosure costs, making public disclosure more sensitive.

We also construct post-move career outcomes to examine whether pre-move public activity is associated with labor-market consequences. These outcomes include compensation changes, indicators for compensation increases, movement to larger firms, tenure at the destination job, and seniority changes. These measures allow us to examine whether public signaling activity is associated with economically meaningful career outcomes after mobility.

3.6 Firm-Level Sample and Measures

For the firm-level analysis, we aggregate engineer disclosure to the firm-year level. We construct a measure of engineer disclosure intensity based on the share of a firm's engineering workforce with above-median public GitHub activity in a given year (denoted ShareSignaling in the firm-level tables). This measure captures the extent to which the firm's engineers make their skills and technical output publicly visible.

We link this firm-year measure to firm-level financial outcomes. The main firm-level outcomes are earnings persistence. Earnings persistence captures the extent to which current

performance changes predict future performance.[1] The firm-level outcome allows us to examine whether public engineer disclosure is associated with knowledge redistribution at the firm level.

3.7 Descriptive Statistics

Table 1 reports descriptive statistics for the sample. Panel A describes the event panel, including the number of engineers, transitions, and engineer-day observations. Panel B reports daily GitHub activity measures. Public GitHub activity is sparse at the daily level: most engineer-days have zero commits, pull requests, repositories created, forks, active forks, or merged pull requests. This sparsity motivates the use of inverse hyperbolic sine transformations in the regression analyses.

Panel C reports post-move career outcomes, including compensation changes, destination-job tenure, movement to larger firms, and seniority changes. Panel C is computed from the position-level transition cross-section, which is larger than the event-window sample because it does not require construction of the daily GitHub activity panel around each transition. These outcomes provide the basis for examining whether pre-move public activity is associated with subsequent labor-market consequences.

Overall, the linked GitHub–Revelio data provide a setting in which we can observe both the timing of worker mobility and the public production and diffusion of technical knowledge. This structure allows us to test whether labor-market frictions generate public disclosure before job transitions and whether that disclosure is associated with knowledge diffusion, worker outcomes, and firm-level consequences.

[1] Another outcome variable is employment in Online Appendix Table A3. Headcount momentum captures whether current employment growth continues into subsequent periods.

4. Empirical Design

4.1 Event-Study Design Around Job Transitions

We begin by examining how engineers' public GitHub activity evolves around employer-to-employer transitions. The event-study design allows us to test whether public coding activity increases before job transitions, whether it declines after the transition, and whether the timing of the change is consistent with pre-move labor-market signaling.

For each job transition, we define event time relative to the start date of the engineer's new job. We construct an engineer-by-day panel over the 360 days before and after the transition. Our baseline specification estimates changes in GitHub activity across event-time windows:

$$Activity_{i,d,e} = \alpha_e + \gamma_d + \sum_m \delta_m \, Window_{m,e,d} + \varepsilon_{i,d,e}.$$

Here, $Activity_{i,d,e}$ is a measure of public GitHub activity for engineer i on calendar day d around transition event e. The term α_e denotes event fixed effects, which absorb time-invariant differences across engineer-transition events, including stable differences in an engineer's baseline GitHub activity, ability, and employer characteristics at the time of the move. The term γ_d denotes calendar-day fixed effects, which absorb platform-wide activity trends, seasonality, macroeconomic conditions, and other shocks common to all GitHub users on a given day.

The indicators $Window_{m,e,d}$ capture event-time windows relative to the transition date. In the main specification, we compare activity in the near-pre window [-180,-1], near-post window [0,+180], and far-post window [+181,+360] to the far-pre window [-360,-181]. We also estimate monthly event-time specifications to trace the dynamic pattern of activity before and after the move. Standard errors are clustered at the engineer level.

This design identifies within-event changes in activity relative to a baseline period far from the transition. The key empirical pattern of interest is whether activity rises before the move and then declines after the move. Such a pattern is difficult to reconcile with a permanent increase in developer productivity or a general trend in GitHub adoption. Instead, it is consistent with engineers increasing public activity when the labor-market value of signaling is highest.

We estimate this specification separately for two sets of outcomes. The first set captures public signaling activity: commits, pull requests opened, and repositories created. The second set captures knowledge diffusion: forks received, active forks, and pull requests merged. All activity measures are transformed using the inverse hyperbolic sine transformation to accommodate the large mass of zero observations while preserving variation among active engineer-days.

4.2 Cross-Sectional Tests of Signaling Incentives and Proprietary Costs

The event-study design establishes whether public activity changes around job transitions. We next examine whether these changes vary in ways predicted by the signaling and disclosure cost mechanisms. Specifically, we test whether the pre-move increase is attenuated when workers already possess stronger labor-market signals or when the cost of disclosure is higher.

We estimate specifications of the following form:

$$Activity_{i,d,e} = \alpha_e + \gamma_d + \beta PreMove_{e,d} + \theta(PreMove_{e,d} \times Z_{i,e}) + \varepsilon_{i,d,e}.$$

Here, $PreMove_{e,d}$ is an indicator for the pre-transition window, and $Z_{i,e}$ is a worker, firm, or market characteristic measured before the transition. The coefficient θ captures whether the pre-move increase in public GitHub activity differs across groups.

We consider several dimensions of heterogeneity. First, we examine worker seniority and compensation. Senior and higher-compensation engineers likely possess stronger pre-existing labor-market signals, reducing the marginal value of additional public disclosure. Second, we examine R&D intensity at the origin firm. Engineers at R&D-intensive firms may face greater proprietary-cost concerns, making public disclosure more sensitive. Third, we examine labor-market conditions, including the COVID period, when demand for software engineers and the normalization of remote work may have reduced the need for additional signaling.

These tests are useful for distinguishing the signaling mechanism from alternative explanations. If pre-move GitHub activity simply reflected general changes in workload, job preparation, or platform usage, it is less clear why the response should be systematically weaker among engineers with stronger credentials or at firms with greater proprietary-cost concerns. Cross-sectional variation therefore provides mechanism evidence that complements the timing evidence from the event study.

4.3 Worker-Level Career Outcomes

We next examine whether pre-move public activity is associated with subsequent labor-market outcomes. This analysis asks whether the activity we interpret as signaling is valued in the labor market.

We aggregate GitHub activity over the pre-move window and estimate regressions of post-move outcomes on pre-move activity:

$$Outcome_{i,e}^{post} = \alpha + \beta PreMoveActivity_{i,e} + X_{i,e}'\Gamma + \eta_t + \varepsilon_{i,e}.$$

The outcome variables include compensation changes, an indicator for compensation increases, movement to larger firms, destination-job tenure, seniority changes, and indicators for upward mobility. $PreMoveActivity_{i,e}$ captures the engineer's public GitHub activity in the pre-transition period. In some specifications, we distinguish between general public activity and externally validated activity, such as forks received or merged pull requests. The vector $X_{i,e}$ includes worker and origin-job characteristics, such as seniority, compensation, origin-firm size, and other controls available in Revelio and Compustat. We include time fixed effects to absorb common labor-market conditions.

This analysis is not designed to establish that GitHub activity causally improves career outcomes. Engineers who increase public activity may differ from those who do not in unobserved ability, ambition, or job-search intensity. Instead, the purpose is to test whether pre-move public disclosure is associated with economically meaningful labor-market outcomes. A positive association is consistent with the interpretation that public GitHub activity contains information valued by employers, particularly when the activity receives external validation from other users.

4.4 Instrumental Variable Design Using Firm-Level Headcount Contractions

A central concern in the baseline event-study analysis is selection. Engineers who are more ambitious, more able, or more likely to change jobs may also be more likely to increase GitHub activity. To address this concern, we use firm-level headcount contractions as an instrument for worker departure.

The identifying idea is that contractions at the origin firm increase the probability that some engineers leave for reasons related to firm-level labor demand rather than individual signaling

propensity. These contraction-induced departures provide variation in mobility that is less likely to be driven by the engineer's unobserved career ambition or latent GitHub activity.

The first-stage specification estimates whether firm-level headcount contractions predict individual departure:

$$Leaver_{i,f,t} = \alpha + \pi HeadcountContraction_{f,t} + X'_{i,f,t}\Gamma + \eta_t + \varepsilon_{i,f,t}.$$

Here, $Leaver_{i,f,t}$ is an indicator for whether engineer i leaves firm f following the contraction period. $HeadcountContraction_{f,t}$ is measured using firm-level employment declines. We use both a continuous measure of headcount contraction and binary indicators for severe contractions, such as contractions exceeding specified thresholds. Controls include seniority, firm size, and shock-year fixed effects. Standard errors are clustered at the firm level.

The second-stage specification estimates the effect of instrumented departure on changes in public GitHub activity:

$$\Delta Activity_i = \alpha + \beta \widehat{Leaver}_i + X'_i\Gamma + \eta_t + \varepsilon_i.$$

The dependent variable is the change in activity from the pre-period to the post-period around the firm-level shock. The coefficient β captures the local average treatment effect of departure for engineers whose mobility is induced by firm-level headcount contractions.

This design helps address the concern that the baseline relation between mobility and public GitHub activity is driven by self-selection into departure. If contraction-induced departures lead to increases in public coding activity, the evidence supports the interpretation that mobility incentives cause workers to increase public skill demonstration.

We also examine dose-response patterns by varying the severity threshold used to define headcount contractions. If the mechanism is driven by labor-market shocks that increase departure risk, stronger contractions should generate stronger first-stage effects and larger activity responses. These analyses help validate the interpretation that the instrument operates through induced mobility rather than unrelated firm-level changes. The exclusion restriction requires that firm-level headcount contractions affect engineers' GitHub activity primarily through their effect on departure incentives.

4.5 Firm-Level Consequences

The final set of analyses examines whether worker-level disclosure is associated with firm-level outcomes. The theoretical framework suggests that public disclosure may redistribute knowledge outside the firm, weakening persistence in firm performance.

We construct a firm-year measure of engineer disclosure intensity, $ShareSignaling_{f,t}$, defined as the share of a firm's engineering workforce with above-median public GitHub activity in year t. We then estimate firm-level panel regressions of the following form:

$$Outcome_{f,t+k} = \alpha + \beta_1 ShareSignaling_{f,t} + \beta_2 \Delta Outcome_{f,t} + \beta_3 (ShareSignaling_{f,t} \times \Delta Outcome_{f,t}) + \theta_f + \psi_t + \varepsilon_{f,t}.$$

Here, $Outcome_{f,t+k}$ is a future firm-level outcome measured k years after the signaling year. We examine earnings-related outcomes, such as ROA and operating performance. The term $\Delta Outcome_{f,t}$ captures contemporaneous changes in firm performance. The interaction coefficient β_3 measures whether the persistence of the outcome differs when engineer disclosure is high.

For earnings outcomes, a negative interaction between engineer disclosure and current earnings changes is consistent with lower earnings persistence when more engineers publicly disclose. This pattern is consistent with knowledge redistribution or reduced persistence of private human-capital advantages.

The firm-level analyses are designed to provide evidence on broader economic implications rather than definitive causal estimates. We include firm fixed effects or industry-by-year fixed effects, depending on the specification, to absorb persistent firm differences and time-varying industry shocks.

5. Results

5.1 Public GitHub Activity Around Job Transitions

We begin by examining whether engineers increase public GitHub activity anticipating their potential job transitions. Table 2 reports event-study estimates for three measures of public signaling activity: commits, pull requests opened, and repositories created. These measures capture different dimensions of publicly observable output, ranging from general coding activity to more deliberate forms of public project creation.

The results show an overall increase in public activity in the months preceding a job transition. Relative to an earlier baseline period, engineers significantly increase commits and repository creation in the months just before the move: pull requests opened rise as well, though this increase is statistically significant only in the pooled pre/post comparison (Table 2, Panel B). The monthly estimates indicate that the rise begins several months before the move and peaks shortly before the transition, consistent with engineers increasing public output when the labor-market value of signaling is highest.

The increase is economically meaningful given the sparse, intermittent nature of daily GitHub activity and the large share of engineer-days with zero activity. Extensive-margin estimates using indicator measures of daily activity (reported in Online Appendix Table A1) imply that, in the near-pre window, the daily probability of any public commit rises by roughly 9 percent and the probability of creating a new public repository by roughly 13 percent, relative to each engineer's far pre-move baseline. In level terms, over the 180-day near-pre window each transition is associated with approximately 0.72 additional days with public coding activity and 0.31 additional days on which a new public repository was created.

Activity declines sharply after the transition. Commits, pull requests opened, and repositories created all fall in the post-move period, significantly below the far pre-move baseline (Table 2, Panel A). This post-move decline is what distinguishes the signaling interpretation from alternatives such as a persistent productivity gain, general career progression, or secular growth in GitHub use: were engineers simply becoming more productive or more attached to the platform, we would not expect activity to fall immediately after the move. The reversal indicates instead that public output is temporarily elevated around the job-search period.

Figure 1 plots the monthly event-study coefficients for commits around job transitions. Commits begin to rise several months before the transition, peak shortly before the move, and decline after the new job begins. This dynamic pattern complements the grouped-window estimates in Table 2. Taken together, Table 2 provides the first piece of evidence that public GitHub activity is systematically related to labor-market transitions: engineers increase visible technical output before moving and reduce it afterward.

5.2 Knowledge Diffusion Around Job Transitions

Table 3 examines whether pre-move public activity is accompanied by knowledge diffusion. We measure diffusion using external engagement with the engineer's public work: forks received, active forks, and pull requests merged into the engineer's repositories. These measures differ in how much engagement they require: forks received capture passive copying of a repository, whereas active forks (copies that are subsequently developed) and merged pull requests (outside contributions integrated back into the engineer's repository) require others to do something beyond a one-click copy.

This distinction is important because public disclosure does not necessarily imply knowledge spillovers. A worker may post code that no one uses. By examining external engagement, we can assess whether public output is taken up by other developers. Forks and active forks indicate that outside users copy or build on the engineer's repositories, while merged pull requests indicate collaborative integration of outside contributions.

The evidence here is weaker and more qualified than the signaling evidence in Table 2. In the near-pre window, none of the three diffusion measures differs significantly from baseline (Panel A). Statistically significant pre-move differences emerge only in the pooled pre/post specification (Panel B), and only for active forks ($p < 0.05$) and merged pull requests ($p < 0.01$); forks received, passive copying that requires no further engagement, is insignificant throughout. For merged pull requests, moreover, much of the pre/post difference reflects a decline in external integration after the move rather than a pronounced pre-move surge. We therefore read the diffusion evidence as suggestive and measure-specific, not as a uniform increase in spillovers across external-engagement outcomes.

Figure 2 plots the monthly event-study coefficients for merged pull requests around job transitions, the diffusion measure with the most robust pre/post difference. The figure is consistent

with the grouped-window estimates: external integration shows no pronounced spike in the months just before the move and falls after the new job begins. What it does show, in line with the central argument, is that whatever external engagement the disclosed work attracts is concentrated around the job-search window rather than emerging only after the worker joins the new firm.

The timing evidence, though weaker than for the signaling outcomes, still speaks to the paper's central claim. Traditional labor-mobility spillovers operate only after workers join new firms. The patterns in Table 3 are consistent with some externally validated use of disclosed work arising before the worker changes employers, because workers make output publicly observable while preparing for the labor market and other participants can use it immediately. We are careful not to overstate this: the diffusion results are statistically weaker and more measure-specific than the signaling results, and we treat them as suggestive evidence on the spillover consequences of pre-move disclosure rather than as established fact.

5.3 Cross-Sectional Evidence on Signaling Incentives and Disclosure Costs

Table 4 examines whether the pre-move increase in public activity varies with worker, market, and firm characteristics. These tests assess whether the baseline pattern behaves as a signaling response should: strengthening when the marginal benefit of a public signal is high, and weakening when workers already hold strong credentials or bear greater individual costs of disclosure.

The first set of tests examines worker characteristics. The pre-move increase is markedly stronger among junior engineers and weaker among higher-compensation engineers (Panels A and B). Junior engineers' pre-move rise in commits is roughly two-and-a-half times that of more senior engineers, whereas higher-compensation engineers exhibit only about a third of the commit response of their lower-paid peers. This is what a signaling hypothesis predicts: the marginal value

of additional public output is greatest for workers whose ability is otherwise hard to observe, and smallest for those who can already point to experience, employer pedigree, and compensation history as credible signals.

The second set of tests examines labor-market conditions. We focus on the COVID-19 period, when demand for software engineers surged, which is driven by rapid digitalization, the expansion of remote work, and intensified hiring by technology firms, labor supply adjusted more slowly, producing more recruiter outreach, faster hiring cycles, and stronger outside options for workers. The pre-move increase is significantly weaker during this period (Panel C), with a commit ramp roughly 40 percent smaller than outside the COVID window. This is consistent with workers needing less additional public signaling when outside options are abundant, and it helps separate the signaling mechanism from a mechanical technology-adoption story: had GitHub activity simply risen during the pandemic, we would expect a higher level of activity, not a weaker transition-related ramp.

The third set of tests examines firm characteristics. Engineers at R&D-intensive firms show a weaker pre-move increase (Panel D), with a commit ramp roughly 40 percent below that of engineers at less R&D-intensive firms. We interpret this as evidence that the individual cost of disclosure is higher for these workers: engineers at R&D-intensive firms are more likely to work on sensitive or less readily generalizable technical knowledge, so publicly demonstrating it is personally costlier or more constrained, whether through employer restrictions they internalize or the risk of exposing proprietary work. Workers' signaling incentives are tempered by the costs they themselves bear from making firm knowledge publicly observable.

Overall, the cross-sectional evidence reinforces the event-study interpretation. The pre-move increase is strongest where the benefits of signaling are likely highest among junior

engineers and in looser labor markets and weakest where workers already hold strong credentials or face greater individual costs of disclosure.

5.4 Worker Career Consequences

Table 5 examines whether pre-move public activity is associated with subsequent labor-market outcomes: the change in compensation at the move and the likelihood of moving to a larger firm. These tests ask whether the activity we interpret as signaling is in fact valued in the labor market.

Table 5 reveals a nuanced pattern. Raw signaling measures are not positively associated with career outcomes; if anything, the associations are negative. Repository creation is significantly negatively associated with compensation gains and, more weakly, with moves to larger firms, and pull requests opened are negatively associated with compensation gains, while commits are unrelated to either outcome (Panel A). In contrast, externally validated activity, especially forks received, is positively and significantly associated with both higher compensation and moves to larger firms (Panel B). The labor market appears to reward public output that other developers take up, rather than the sheer volume of activity a worker generates.

This pattern is consistent with a labor-market signaling interpretation in which validated signals may carry more weight. Public GitHub activity provides prospective employers with verifiable evidence of technical ability, specialization, and engagement with the developer community; but activity that no one engages with may be a weak or noisy signal. Externally validated activity is plausibly more informative because it reflects not only the worker's decision to disclose but also the market's response to that disclosure.

At the same time, we interpret these results with caution. Engineers who increase public GitHub activity before moving may differ from other engineers in ambition, ability, or job-search intensity, so we do not read the evidence as establishing that GitHub activity leads to better labor-market outcomes. Rather, the results indicate that pre-move public activity is economically meaningful and associated with outcomes consistent with the labor market valuing credible public signals.

5.5 Instrumental Variable Evidence from Firm-Level Headcount Contractions

Tables 6 and 7 address selection concerns by using firm-level headcount contractions as an instrument for worker departure. The concern is that engineers who increase public activity before moving may differ systematically from those who do not: more ambitious or more able engineers may both signal more and be more likely to leave. The instrument isolates departures induced by employer-side shocks rather than by individual career choices.

The first stage is strong. Headcount contractions significantly increase the probability that an engineer leaves the firm, and the instrument is powerful under both continuous and threshold-based definitions (K-P F-statistics of 35.8 and 25.2, respectively; Table 6). Firm-level employment declines thus generate variation in departure risk that is not driven solely by individual career choice.

The second-stage estimates indicate that contraction-induced departures lead to economically meaningful increases in public GitHub activity, concentrated in the worker-controlled signaling outcomes: repository creation (significant at the 1 percent level), commits, and pull requests opened (Table 7). These are the outcomes most directly under the engineer's control and therefore provide the cleanest evidence of deliberate public skill demonstration. The 2SLS estimates are substantially larger than their OLS counterparts, indicating that among the

compliers whose departures are induced by employer-side contractions, the pre-move public-activity response is far stronger than the average cross-sectional association.

A dose-response analysis reinforces this interpretation. More severe contractions produce stronger first-stage coefficients, and the public-activity response is statistically significant only at the more severe thresholds (contractions of 20 percent or more), with the largest estimates around the 25 percent threshold (Table 7). This gradient is difficult to reconcile with a purely selection-based explanation and is consistent with employer-side shocks raising mobility incentives, which in turn lead workers to make their skills more publicly observable.

5.6 Firm-Level Consequences

Finally, we examine whether worker-level disclosure has firm-level consequences. Public engineer activity can affect the firm in two offsetting ways. It may erode the firm's informational advantages: when engineers publicly demonstrate skills and approaches, outside parties can learn from or build on that knowledge, potentially reducing the persistence of firm-specific performance (knowledge redistribution).

Table 8 provides nuanced but economically interpretable evidence consistent with knowledge redistribution. The interaction between a firm's current earnings change and the disclosure intensity of its engineers is negative, indicating that earnings are less persistent at firms whose engineers disclose more publicly. The effect is concentrated in operating performance, the dependent variable closest to core operations, where it is statistically significant in the full sample and substantially larger among firms with rising earnings (Panel A). It is also pronounced in the decomposed measures (Panel B): the largest reductions in persistence appear for externally validated activity, forks received and active forks, and for repository creation, rather than for raw commit volume. This is the pattern a knowledge-spillover mechanism predicts, since it is precisely

the disclosed output that others take up and build on that should erode a firm's informational advantage. However, the effect is not uniform: the interaction is insignificant for net ROA in the full sample, so we read the evidence as consistent with, rather than definitive proof of, a knowledge-redistribution channel. Importantly, the mechanism does not require leakage of proprietary code: even generalizable technical approaches, demonstrations of employee capability, or reusable solutions can weaken the informational advantages embedded in a firm's workforce.

The firm-level tests should be viewed as associative and more preliminary than the worker-level event studies. Greater disclosure is associated with lower earnings persistence in several specifications, which is consistent with public disclosure redistributing knowledge away from the originating firm.

6. Conclusion

Labor mobility is widely viewed as a central channel through which knowledge diffuses across firms, with the transfer occurring when workers join a new employer. This paper shows that the diffusion can begin earlier—before workers move at all. When labor markets are characterized by information frictions, job-seeking workers have incentives to make their skills publicly observable, and in software engineering they can do so directly through GitHub. Public coding activity helps workers demonstrate ability to prospective employers, but the same activity makes technical knowledge available to a far broader audience. Career concerns that are entirely private in motivation can therefore generate public knowledge spillovers.

Using our linked sample of GitHub activity and employment histories, we document that engineers increase public GitHub activity in the months before a transition, peaking shortly before the move and declining sharply afterward. This timing is difficult to reconcile with rising

productivity or general platform adoption, which would not reverse at the moment of the move. The increase is weaker precisely where signaling theory predicts: among senior and higher-paid engineers, who already hold strong credentials; during the tight COVID-era labor market, when outside options are easier to obtain; and at R&D-intensive firms, where disclosure is more costly. We find more qualified evidence that this public activity attracts external engagement from other developers, indicating that disclosure can translate into knowledge diffusion before the worker changes firms.

The pre-move disclosure is also economically meaningful. Externally validated activity, especially forks received, is associated with larger compensation gains and moves to larger firms, whereas raw signaling volume is, if anything, negatively related to these outcomes, suggesting that the labor market rewards public work that others validate rather than visible effort per se. To address the concern that engineers who disclose differ systematically from those who do not, we use firm-level headcount contractions as an instrument for departure and find that contraction-induced departures raise worker-controlled public coding activity, with the response strengthening as contractions become more severe. This evidence supports the interpretation that mobility incentives cause workers to increase public skill demonstration. Finally, firms with greater engineer disclosure are less likely to sustain positive earnings changes, consistent with a proprietary cost: disclosed knowledge can be learned from and reused outside the firm, eroding the persistence of its competitive advantage.

These findings contribute to three literatures. To work on labor mobility and knowledge diffusion, we show that spillovers need not wait for workers to arrive at a new employer; the anticipation of mobility can itself induce knowledge disclosure. To work on labor-market signaling, we show that signals can carry real externalities: public coding activity is not merely informative

about the worker but is a productive artifact that others can fork, extend, and build on. And to the accounting and disclosure literature, we identify a worker-level disclosure channel that shapes both labor-market matching and firm-level outcomes, extending disclosure analysis beyond the firm as the unit of disclosure to the individual employee. The unifying insight is that labor-market frictions determine not only who moves across firms, but what knowledge becomes publicly available before the move.

These findings carry implications beyond the workers themselves. For firms, employee disclosure creates a tension between the private benefits workers capture—better matches and stronger outside options—and a cost the firm bears, as publicly disclosed knowledge erodes the persistence of its competitive advantage. For policymakers and platform designers, the results point to an underappreciated source of public-good provision: private career incentives can lead workers to produce publicly useful knowledge even when no one intends it. Several caveats temper these conclusions. We observe only public GitHub activity, not private or internal work, so our estimates speak to disclosed rather than total knowledge production; some pre-move activity may reflect learning or portfolio-building rather than signaling; and our firm-level results are associational. Even so, the central pattern is clear: as more work becomes publicly observable through online platforms, the boundary between labor-market signaling and knowledge disclosure blurs, and the boundaries of the firm are increasingly shaped by the digital signals its workers send.

References

Abou El-Komboz, L., and M. Goldbeck. 2025. Career concerns as a public good: The role of signaling for open source software development. Labour Economics 97: 102800.

Agrawal, A., N. Lacetera, and E. Lyons. 2016. Does standardized information in online markets disproportionately benefit job applicants from less developed countries? Journal of International Economics 103: 1–12.

Almeida, P., and B. Kogut. 1999. Localization of knowledge and the mobility of engineers in regional networks. Management Science 45 (7): 905–917.

Arrow, K. J. 1962. The economic implications of learning by doing. Review of Economic Studies 29 (3): 155–173.

Barrios, J. M., and J. Gallemore. 2024. Tax planning knowledge diffusion via the labor market. Management Science 70 (2): 1194–1215.

Becker, G. S. 1964. Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education. New York, NY: Columbia University Press.

Burks, Stephen V., Bo Cowgill, Michell Hoffman, and Michael Housman. 2015. The Value of Hiring Through Employee Referrals. The Quarterly Journal of Economics 130: 805–39.

Chevalier, J., and G. Ellison. 1999. Career concerns of mutual fund managers. Quarterly Journal of Economics 114 (2): 389–432.

Darrough, M. N., and N. M. Stoughton. 1990. Financial disclosure policy in an entry game. Journal of Accounting and Economics 12 (1–3): 219–243.

Dye, R. A. 1985. Disclosure of nonproprietary information. Journal of Accounting Research 23 (1): 123–145.

Fershtman, C., and N. Gandal. 2011. Direct and indirect knowledge spillovers: The "social network" of open-source projects. RAND Journal of Economics 42 (1): 70–91.

Gow, I. D., A. S. Wahid, and G. Yu. 2018. Managing reputation: Evidence from biographies of corporate directors. Journal of Accounting and Economics 66 (2–3): 448–469.

Greenwald, B. C. 1986. Adverse selection in the labour market. Review of Economic Studies 53 (3): 325–347.

Holmström, B. 1999. Managerial incentive problems: A dynamic perspective. Review of Economic Studies 66 (1): 169–182.

Hung, Chung-Yu, Anne M. Lillis, and Anne Wu. 2023. Does Referral-Based Hiring Exacerbate Agency Problems? The Accounting Review 98 (2): 277–97.

Jaffe, A. B., M. Trajtenberg, and R. Henderson. 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. Quarterly Journal of Economics 108 (3): 577–598.

Kim, J., and K. Valentine. 2021. The innovation consequences of mandatory patent disclosures. Journal of Accounting and Economics 71 (2–3): 101381.

Lerner, J., and J. Tirole. 2002. Some simple economics of open source. Journal of Industrial Economics 50 (2): 197–234.

Marlow, J., and L. Dabbish. 2013. Activity traces and signals in software developer recruitment and hiring. In Proceedings of the 2013 Conference on Computer Supported Cooperative Work, 145–156.

McDermott, G. R., and B. Hansen. 2021. Labor reallocation and remote work during COVID-19: Real-time evidence from GitHub. NBER Working Paper No. 29598.

Pallais, A. 2014. Inefficient hiring in entry-level labor markets. American Economic Review 104 (11): 3565–3599.

Rosenkopf, L., and P. Almeida. 2003. Overcoming local search through alliances and mobility. Management Science 49 (6): 751–766.

Song, J., P. Almeida, and G. Wu. 2003. Learning-by-hiring: When is mobility more likely to facilitate interfirm knowledge transfer? Management Science 49 (4): 351–365.

Spence, M. 1973. Job market signaling. Quarterly Journal of Economics 87 (3): 355–374.

Tambe, P. and Hitt, L.M. 2014. Job hopping, information technology spillovers, and productivity growth. Management science. 60 (2): 338–355.

Verrecchia, R. E. 1983. Discretionary disclosure. Journal of Accounting and Economics 5: 179–194.

Zander, U., and B. Kogut. 1995. Knowledge and the speed of the transfer and imitation of organizational capabilities: An empirical test. Organization Science 6 (1): 76–92.

Table 1: Descriptive Statistics

This table reports descriptive statistics for the sample. Panel A summarizes the engineer-day event-window sample used in the event-study analyses (numbers of engineers, transitions, and engineer-day observations). Panel B reports the mean, dispersion, and pre- versus post-move averages of the six daily GitHub activity measures; in Panel B, ***, **, and * denote significance of the post−pre difference at the 1%, 5%, and 10% levels. Panel C reports engineer and origin-firm characteristics measured at the first-move event level. Please refer to Appendix Table A4 for detailed variable definitions.

Panel A: Sample Overview

Total observations (engineer × day)	24,936,731
Unique engineers	22,293
Unique job transitions in event-window sample (events)	46,637
Sample date range	January 2011 – March 2021
Event window (days relative to move)	−360 to +360
Events per engineer: mean (SD)	2.09 (1.32)
Events per engineer: median [p25, p75]	2 [1, 3]
Days observed per event: mean (SD)	534.7 (275.7)

Panel B: Daily GitHub Activity (N = 24,936,731 engineer-days)

	Mean	SD	P25	P50	% Zero	Pre Mean	Post Mean	Δ(Post−Pre)
Commits	0.185	3.468	0.000	0.000	96.2	0.217	0.153	−0.064***
PRs opened	0.012	0.504	0.000	0.000	99.4	0.013	0.011	−0.003***
Repos created	0.016	0.411	0.000	0.000	98.9	0.019	0.012	−0.007***
Forks received	0.005	0.131	0.000	0.000	99.7	0.005	0.005	−0.000
Active forks	0.001	0.034	0.000	0.000	99.9	0.001	0.001	−0.000***
PRs merged	0.006	0.405	0.000	0.000	99.7	0.007	0.006	−0.001***

Panel C: Engineer and Origin-Firm Characteristics (First Move, Event Level)

	N	Mean	SD	P25	Median	P75
Continuous variables						
Seniority (1–7 scale)	22,190	2.61	1.26	2.00	2.00	3.00
Total compensation ($000s)	22,185	136.7	75.3	86.6	123.8	168.6
Employees (000s)	18,417	159.4	321.4	8.4	47.6	161.4
Total assets ($M)	19,041	138,571	378,923	4,645	31,342	131,310
R&D / Sales	14,446	0.088	0.082	0.031	0.073	0.125
Binary indicators (share = 1)						
High R&D intensity	14,446	0.500				
High compensation	22,185	0.500				

Table 2: Mobility and Signaling Activity

This table reports event-study estimates of public signaling activity around employer-to-employer job transitions. The dependent variables are the inverse hyperbolic sine (IHS) of daily commits, pull requests opened, and repositories created. Panel A reports coefficients on grouped event-time windows relative to the far-pre window [−360, −181]; Panel B reports the pooled pre/post specification, reporting the pre-move window [−360, 0] relative to the omitted post-move window. Please refer to Appendix Table A4 for detailed variable definitions. All specifications include event and calendar-day fixed effects. t-statistics based on standard errors clustered at the engineer level are reported in parentheses. ***, **, * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Grouped Event Windows

	IHS(Commits)	*IHS(PRs Opened)*	*IHS(Repos Created)*
	(1)	*(2)*	*(3)*
Near pre [−180, −1]	0.0069***	0.0003	0.0019***
	(6.6433)	(0.8466)	(7.8942)
Near post [0, +180]	−0.0216***	−0.0016***	−0.0014***
	(−14.0532)	(−3.9583)	(−4.1634)
Far post [+181, +360]	−0.0227***	−0.0014***	−0.0013***
	(−11.4806)	(−2.9919)	(−2.9284)
Observations	24,936,197	24,936,197	24,936,197
N Engineers	22,154	22,154	22,154
R^2	0.147	0.117	0.040
Adj. R^2	0.145	0.115	0.038
Event-ID FE	Yes	Yes	Yes
Calendar-date FE	Yes	Yes	Yes

Panel B: Pre/Post Specification

	IHS(Commits)	*IHS(PRs Opened)*	*IHS(Repos Created)*
	(1)	*(2)*	*(3)*
Pre-move [−360, 0]	0.0300***	0.0019***	0.0038***
	(24.4927)	(6.9749)	(16.1545)
Observations	24,936,197	24,936,197	24,936,197
N Engineers	22,154	22,154	22,154
R^2	0.146	0.117	0.040
Adj. R^2	0.145	0.115	0.038
Event-ID FE	Yes	Yes	Yes
Calendar-date FE	Yes	Yes	Yes

Table 3: Knowledge Diffusion Around Job Transitions

This table reports event-study estimates of knowledge diffusion around employer-to-employer job transitions. The dependent variables are the inverse hyperbolic sine (IHS) of daily forks received, active forks, and pull requests merged into the engineer's repositories. Panel A reports coefficients on grouped event-time windows relative to the far-pre window [−360, −181]; Panel B reports the pooled pre/post specification, reporting the pre-move window [−360, 0] relative to the omitted post-move window. Please refer to Appendix Table A4 for detailed variable definitions. All specifications include event and calendar-day fixed effects. t-statistics based on standard errors clustered at the engineer level are reported in parentheses. ***, **, * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Grouped Event Windows

	IHS(Forks Received)	*IHS(Active Forks)*	*IHS(PRs Merged)*
	(1)	*(2)*	*(3)*
Near pre [−180, −1]	0.0001	0.0000	0.0002
	(1.2484)	(0.8109)	(1.0421)
Near post [0, +180]	−0.0001	−0.0000	−0.0010***
	(−0.4099)	(−0.7501)	(−3.3619)
Far post [+181, +360]	−0.0003	−0.0000	−0.0011***
	(−1.4053)	(−0.5782)	(−3.1869)
Observations	24,936,197	24,936,197	24,936,197
N Engineers	22,154	22,154	22,154
R^2	0.273	0.023	0.123
Adj. R^2	0.272	0.021	0.121
Event-ID FE	Yes	Yes	Yes
Calendar-date FE	Yes	Yes	Yes

Panel B: Pre/Post Specification

	IHS(Forks Received)	*IHS(Active Forks)*	*IHS(PRs Merged)*
	(1)	*(2)*	*(3)*
Pre-move [−360, 0]	0.0002	0.0001**	0.0012***
	(1.4993)	(2.2183)	(7.0303)
Observations	24,936,197	24,936,197	24,936,197
N Engineers	22,154	22,154	22,154
R^2	0.273	0.023	0.123
Adj. R^2	0.272	0.021	0.121
Event-ID FE	Yes	Yes	Yes
Calendar-date FE	Yes	Yes	Yes

Table 4: Cross-Sectional Heterogeneity in Pre-Move Signaling and Knowledge Diffusion

This table reports interaction estimates of how the pre-move increase in public GitHub activity varies with engineer and firm characteristics. The dependent variables are the IHS of the six daily GitHub measures (columns), and PreMove is an indicator for engineer-days in the pre-move window. Panel A interacts PreMove with Junior (origin seniority ≤ 3, i.e., below Manager); Panel B with High Compensation; Panel C with the COVID-19 period (March 2020–December 2021); and Panel D with High R&D intensity. Please refer to Appendix Table A4 for detailed variable definitions. All specifications include event and calendar-day fixed effects. t-statistics based on standard errors clustered at the engineer level are reported in parentheses. ***, **, * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Junior Engineers

		Signaling		*Knowledge diffusion*		
	IHS(Commits)	*IHS(PRs Opened)*	*IHS(Repos Created)*	*IHS(Forks Recv.)*	*IHS(Active Forks)*	*IHS(PRs Merged)*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
PreMove	0.0093***	−0.0011	0.0001	−0.0006	−0.0001*	−0.0009
	(3.1764)	(−0.9926)	(0.1813)	(−1.3913)	(−1.8012)	(−0.8383)
PreMove × Junior	0.0143***	0.0028**	0.0030***	0.0011**	0.0002***	0.0020*
	(4.4348)	(2.1560)	(5.1898)	(2.2274)	(3.1269)	(1.7094)
Observations	12,143,649	12,143,649	12,143,649	12,143,649	12,143,649	12,143,649
R^2	0.1461	0.1507	0.0384	0.3586	0.0284	0.1849
Adj. R^2	0.1443	0.1489	0.0363	0.3573	0.0264	0.1832

Panel B: High Compensation

		Signaling		*Knowledge diffusion*		
	IHS(Commits)	*IHS(PRs Opened)*	*IHS(Repos Created)*	*IHS(Forks Recv.)*	*IHS(Active Forks)*	*IHS(PRs Merged)*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
PreMove	0.0291***	0.0020***	0.0045***	0.0002	0.0001*	0.0012***
	(14.6671)	(4.4234)	(11.4837)	(1.0645)	(1.7305)	(3.6474)
PreMove × High Comp	−0.0185***	−0.0021***	−0.0044***	0.0001	−0.0000	−0.0012**
	(−7.3675)	(−2.9956)	(−9.5902)	(0.2278)	(−0.1984)	(−2.1619)
Observations	12,139,911	12,139,911	12,139,911	12,139,911	12,139,911	12,139,911
R^2	0.1462	0.1508	0.0384	0.3586	0.0284	0.1849
Adj. R^2	0.1444	0.1490	0.0364	0.3573	0.0264	0.1832

Panel C: COVID-19 Period

	Signaling			*Knowledge diffusion*		
	IHS(Commits)	*IHS(PRs Opened)*	*IHS(Repos Created)*	*IHS(Forks Recv.)*	*IHS(Active Forks)*	*IHS(PRs Merged)*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
PreMove	0.0327***	0.0021***	0.0041***	0.0002*	0.0001**	0.0015***
	(24.4186)	(7.0660)	(16.2129)	(1.9533)	(2.2965)	(7.3540)
PreMove × COVID	−0.0123***	−0.0010	−0.0015***	−0.0003	−0.0000	−0.0010**
	(−4.7379)	(−1.4100)	(−3.5001)	(−1.1166)	(−0.4535)	(−2.3133)
Observations	24,936,197	24,936,197	24,936,197	24,936,197	24,936,197	24,936,197
R^2	0.1465	0.1172	0.0401	0.2730	0.0227	0.1229
Adj. R^2	0.1448	0.1154	0.0382	0.2715	0.0208	0.1211

Panel D: High R&D Intensity

	Signaling			*Knowledge diffusion*		
	IHS(Commits)	*IHS(PRs Opened)*	*IHS(Repos Created)*	*IHS(Forks Recv.)*	*IHS(Active Forks)*	*IHS(PRs Merged)*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
PreMove	0.0255***	0.0021***	0.0030***	0.0003	0.0001	0.0012**
	(9.7197)	(3.1013)	(6.3607)	(0.8946)	(1.2129)	(2.4385)
PreMove × High R&D	−0.0104***	−0.0022**	−0.0018***	−0.0001	0.0001	−0.0010
	(−3.1625)	(−2.1939)	(−3.2848)	(−0.2326)	(1.0395)	(−1.2336)
Observations	7,880,693	7,880,693	7,880,693	7,880,693	7,880,693	7,880,693
R^2	0.1600	0.1819	0.0360	0.3936	0.0337	0.2259
Adj. R^2	0.1581	0.1800	0.0338	0.3922	0.0315	0.2241

Table 5: Worker Career Consequences of Pre-Move GitHub Activity

This table reports the association between pre-move GitHub activity and post-move career outcomes; each cell is the coefficient on the indicated pre-move measure from a separate OLS regression. The dependent variables are the log compensation change at the move (Column 1) and an indicator for moving to a larger firm by assets (Column 2). Panel A uses the three signaling measures and Panel B the three knowledge-diffusion measures. All regressions include origin-firm fixed effects and control for seniority, total compensation, log assets, and log employees. Please refer to Appendix Table A4 for detailed variable definitions. t-statistics based on standard errors clustered at the engineer level are reported in parentheses. ***, **, * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Signaling Measures (+ Origin Firm FE)

	Log Δ Comp (at move)	*Larger Firm (by Assets)*
	(1)	(2)
Pre-move avg IHS(Commits)	−0.0239	0.0016
	(−1.561)	(0.092)
Pre-move avg IHS(Repos Created)	−0.2356***	−0.1461*
	(−3.552)	(−1.803)
Pre-move avg IHS(PRs Opened)	−0.0867*	0.0151
	(−1.871)	(0.243)
Observations	23,428	23,286
R^2	0.320	0.196
Adj. R^2	0.270	0.136

Panel B: Knowledge Diffusion Measures (+ Origin Firm FE)

	Log Δ Comp (at move)	*Larger Firm (by Assets)*
	(1)	(2)
Pre-move avg IHS(Forks Received)	0.2120***	0.1950***
	(2.949)	(2.587)
Pre-move avg IHS(Active Forks)	1.5399*	0.1195
	(1.917)	(0.130)
Pre-move avg IHS(PRs Merged)	−0.0895	0.0727
	(−1.524)	(1.288)
Observations	23,428	23,286
R^2	0.321	0.196
Adj. R^2	0.270	0.136

Table 6: Firm-Level Instrumental Variable: Push for Leavers (2SLS)

This table reports instrumental-variable estimates that use firm-level headcount contraction as an instrument for worker departure (Leaver), estimated on the position-level IV sample. Panel A reports first-stage diagnostics for the continuous and binary (≥ 25%) instruments; Panel B reports cross-sectional 2SLS estimates by GitHub outcome. All specifications include seniority and shock-year fixed effects and control for log firm size. Please refer to Appendix Table A4 for detailed variable definitions. t-statistics based on standard errors clustered at the firm level are reported in parentheses. ***, **, * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: First-Stage Diagnostics

Instrument:	*Continuous*	*Binary (≥25%)*
Dep. var.:	*Leaver*	*Leaver*
	(1)	(2)
Headcount contraction	0.2035***	0.0908***
	(5.9821)	(5.0232)
log(firm size)	-0.003	-0.0018
	(-1.4138)	(-0.8977)
K-P F-statistic	35.79	25.23
N	54,015	48,431
Seniority FE	Yes	Yes
Shock-year FE	Yes	Yes
R-sq	0.0716	0.0723
Adj. R-sq	0.0712	0.0719

Panel B: Cross-Sectional 2SLS

	(1)	(2)	(3)
IHS(Repos Created)	0.0009***	0.0061*	0.0123***
	(4.6285)	(1.7558)	(2.6678)
IHS(Commits)	0.0023**	0.0159	0.0377*
	(2.2222)	(0.6591)	(1.6793)
IHS(PRs Opened)	-0.0000	0.0031	0.0058*
	(-0.0070)	(0.6659)	(1.7334)
IHS(PRs Merged)	0.0000	0.0012	0.0018
	(0.0310)	(0.6052)	(1.2914)
IHS(Forks Received)	0.0001	-0.0002	-0.0012
	(1.2838)	(-0.1632)	(-1.1142)
IHS(Active Forks)	0.0000	0.0003	0.0001
	(0.0650)	(1.0987)	(0.3453)
N	54,015	54,015	48,431
log firm size	Yes	Yes	Yes
Seniority FE	Yes	Yes	Yes
Shock-year FE	Yes	Yes	Yes

Table 7: Heterogeneous Effects of the Instrument — 2SLS LATE by Layoff Intensity

This table reports two-stage least squares (2SLS) local-average-treatment-effect (LATE) estimates of worker departure (Leaver) on engineer GitHub activity, instrumenting Leaver with an indicator that the firm's year-over-year headcount declined by at least the indicated threshold (−10% to −30%, computed from Compustat employment). The outcome is the change in IHS-transformed daily activity between the post-move [0, +12] and pre-move [−12, −1] windows. Panel A reports first-stage diagnostics across thresholds; Panel B traces the LATE on repositories created across thresholds; Panel C reports the LATE for all six GitHub outcomes at the −25% threshold, alongside OLS benchmarks from the full cross-section. Please refer to Appendix Table A4 for detailed variable definitions. All specifications include seniority and shock-year fixed effects. t-statistics based on standard errors clustered at the firm level are reported in parentheses. ***, **, * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: First-Stage Diagnostics by Layoff Severity Threshold

Dependent variable: Leaver

	(1)	*(2)*	*(3)*	*(4)*	*(5)*
Instrument (Binary):	*≥ 10%*	*≥ 15%*	*≥ 20%*	*≥ 25%)*	*≥ 30%*
Instrument coefficient	0.0406***	0.0508***	0.0571***	0.0908***	0.1369***
	(4.7370)	(4.4771)	(4.0901)	(5.0232)	(6.6358)
log (firm size)	−0.0030	−0.0025	−0.0024	−0.0018	−0.0017
	(−1.3436)	(−1.1424)	(−1.1271)	(−0.8977)	(−0.8439)
K-P F-statistic	22.44	20.04	16.73	25.23	44.03
Observations	54,015	51,426	49,998	48,431	47,835
Seniority FE	Yes	Yes	Yes	Yes	Yes
Shock-year FE	Yes	Yes	Yes	Yes	Yes
R^2	0.0707	0.0711	0.0712	0.0723	0.0731
Adjusted R^2	0.0703	0.0707	0.0708	0.0719	0.0727

Panel B: 2SLS LATE on Repo Created — Dose–Response Across Thresholds

Dependent variable: Δ IHS(Repo Created), post–pre. Endogenous variable: Leaver. Instrument: Headcount decline indicator at indicated threshold.

	(1)	*(2)*	*(3)*	*(4)*	*(5)*
Instrument (Binary):	*≥ 10%*	*≥ 15%*	*≥ 20%*	*≥ 25%*	*≥ 30%*
Leaver (instrumented)	0.002	0.0019	0.0098**	0.0123***	0.0098***
	(0.3918)	(0.3236)	(2.3544)	(2.6678)	(2.6481)
log (firm size)	0.0001	0.0001	0.0001**	0.0001**	0.0001**
	(1.602)	(1.4879)	(2.0331)	(1.9792)	(2.0165)
K-P F-statistic	22.44	20.04	16.73	25.23	44.03
Observations	54,015	51,426	49,998	48,431	47,835
Seniority FE	Yes	Yes	Yes	Yes	Yes
Shock-year FE	Yes	Yes	Yes	Yes	Yes
R^2	0.0209	0.0202	−0.0590	−0.1062	−0.0595
Adjusted R^2	0.0205	0.0198	−0.0595	−0.1066	−0.0599

Panel C: 2SLS LATE for All GitHub Outcomes at the −25% Threshold (with OLS Benchmark)

Dependent variable: Δ IHS(activity), post–pre. Endogenous variable: Leaver. Instrument (2SLS): Headcount decline ≥ 25%.

	Signaling		*Knowledge diffusion*	

	Commits	*PRs Opened*	*Repo Created*	*Forks Received*	*Active Forks*	*PRs Merged*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
OLS benchmark (full sample)						
Leaver	0.0023**	0.0000	0.0009***	0.0001	0.0000	0.0000
	(2.2222)	(−0.0070)	(4.6285)	(1.2838)	(0.0651)	(0.031)
log (firm size)	0.0001	0.0000	0.0001**	0.0000	0.0000	0.0000
	(1.2433)	(1.3774)	(2.024)	(−1.1587)	(0.4531)	(0.3141)
Observations (OLS)	54,015	54,015	54,015	54,015	54,015	54,015
R^2 (OLS)	0.0203	0.0079	0.0222	0.005	0.0032	0.0052
Adjusted R^2 (OLS)	0.0199	0.0075	0.0218	0.0046	0.0028	0.0048
2SLS (instrument: Headcount decline ≥ 25%)						
Leaver (instrumented)	0.0377*	0.0058*	0.0123***	−0.0012	0.0001	0.0018
	(1.6793)	(1.7334)	(2.6678)	(−1.1142)	(0.3453)	(1.2914)
log (firm size)	0.0003	0.0001*	0.0001**	0.0000	0.0000	0.0000
	(1.5766)	(1.7834)	(1.9792)	(−1.2591)	(0.6066)	(0.9322)
K-P F-statistic	25.23	25.23	25.23	25.23	25.23	25.23
Observations (2SLS)	48,431	48,431	48,431	48,431	48,431	48,431
R^2 (2SLS)	−0.0246	−0.0138	−0.1062	0.0023	0.0025	−0.0008
Adjusted R^2 (2SLS)	−0.0251	−0.0142	−0.1066	0.0019	0.002	−0.0013
Seniority FE	Yes	Yes	Yes	Yes	Yes	Yes
Shock-year FE	Yes	Yes	Yes	Yes	Yes	Yes

Table 8: Firm-Level Consequences — Future Earnings Conditional on Current Earnings and Employee GitHub Activity

This table reports OLS estimates relating future firm earnings to the current earnings change, employee GitHub disclosure intensity (ShareSignaling), and their interaction. Panel A reports the full sample and the negative- and positive-earnings-change subsamples, with ROA and operating ROA as dependent variables; Panel B decomposes ShareSignaling into the six individual GitHub measures (one-year horizon, dependent variable ROA). All specifications include industry-by-year fixed effects and control for log total assets, leverage, cash ratio, Tobin's Q, log employees, and lagged ROA. Please refer to Appendix Table A4 for detailed variable definitions. t-statistics based on standard errors clustered at the firm level are reported in parentheses. ***, **, * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Earnings Persistence by Sign of Earnings Change

Sample	*Full Sample*		*ΔEarnings < 0*		*ΔEarnings > 0*	
Dep. var	*ROA*	*Op.ROA*	*ROA*	*Op.ROA*	*ROA*	*Op.ROA*
	(1)	(2)	(3)	(4)	(5)	(6)
ΔEarnings(f,t)	0.4918***	0.5118***	0.2231**	0.1957***	0.9347***	1.0058***
	(6.9185)	(7.1913)	(2.4395)	(3.0674)	(5.2055)	(7.3777)
ShareSignaling(f,t)	0.0037	-0.0045	0.0050	0.0055	0.0098	0.0109
	(0.4938)	(-0.6933)	(0.4567)	(0.5554)	(1.1350)	(0.9810)
ΔEarnings × ShareSignaling	-0.1418	-0.2408**	-0.0101	0.0534	-0.1549	-0.3799***
	(-1.3195)	(-2.2645)	(-0.0558)	(0.3306)	(-0.9370)	(-3.6220)
Log(Total Assets)	0.0028	-0.0099***	0.0038	-0.0077***	0.0043	-0.0082*
	(1.0067)	(-3.2007)	(1.2466)	(-2.7651)	(0.7860)	(-1.9176)
Leverage	0.0037	0.0856***	0.0082	0.1038***	-0.0169	0.0510***
	(0.2483)	(6.4353)	(0.3492)	(5.2439)	(-0.6922)	(2.6188)
Cash/Assets	-0.0734***	-0.1182***	-0.0341	-0.0663**	-0.0917**	-0.1407***
	(-2.7518)	(-4.8243)	(-1.0827)	(-2.0878)	(-2.0385)	(-4.3875)
Tobin's Q	0.0068***	0.0067***	0.0052**	0.0034	0.0044	0.0051***
	(3.8734)	(4.3726)	(2.5154)	(1.6304)	(1.5937)	(2.6599)
Log(Employees)	0.0042	0.0144***	0.0053**	0.0138***	0.0012	0.0108**
	(1.4850)	(4.8449)	(2.1640)	(4.7797)	(0.1896)	(2.5410)
ROA(f,t−1)	0.5671***	0.5716***	0.6178***	0.6328***	0.8600***	0.8940***
	(12.1943)	(11.7939)	(7.0944)	(7.2516)	(6.3225)	(8.2470)
Constant	-0.0259	0.1181***	-0.0562**	0.0776***	-0.0385	0.1033***
	(-1.1291)	(4.7824)	(-2.2052)	(3.2738)	(-0.9545)	(3.0810)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Clustered by	Firm	Firm	Firm	Firm	Firm	Firm
Observations	4149	4095	2020	1988	1992	1970
R^2	0.4730	0.5487	0.5105	0.5517	0.5371	0.6653
Adj. R^2	0.4250	0.5071	0.4433	0.4895	0.4730	0.6184

Panel B: Decomposed by GitHub Measure (k = 1, DV = ROA)

	Signaling			*Knowledge diffusion*		
ShareMeasure	*Commits*	*PRs*	*Repos*	*Forks*	*Act.Forks*	*PRsMrgd*
Dep. var	*ROA*	*ROA*	*ROA*	*ROA*	*ROA*	*ROA*
	(1)	(2)	(3)	(4)	(5)	(6)
ΔEarnings(f,t)	0.4915***	0.4748***	0.4839***	0.4969***	0.4731***	0.4725***
	(7.0028)	(6.5805)	(6.6817)	(7.7710)	(6.8697)	(6.6200)
ΔEarnings × ShareMeasure	-0.1407	-0.1051	-0.2621**	-0.5526**	-0.8939**	-0.0762
	(-1.2690)	(-0.7869)	(-2.1660)	(-2.1857)	(-2.1630)	(-0.5913)

Share: Commits	0.0068					
	(0.8199)					
Share: PRs Opened		0.0012				
		(0.1497)				
Share: Repos Created			0.0040			
			(0.5712)			
Share: Forks Received				0.0041		
				(0.5307)		
Share: Active Forks					0.0132	
					(1.1321)	
Share: PRs Merged						0.0015
						(0.1607)
Log(Total Assets)	0.0027	0.0025	0.0025	0.0021	0.0022	0.0025
	(0.9796)	(0.9188)	(0.9371)	(0.7907)	(0.8240)	(0.9149)
Leverage	0.0041	0.0030	0.0032	-0.0038	0.0031	0.0028
	(0.2744)	(0.1971)	(0.2139)	(-0.2563)	(0.2057)	(0.1879)
Cash/Assets	-0.0732***	-0.0758***	-0.0763***	-0.0788***	-0.0676**	-0.0762***
	(-2.7397)	(-2.6882)	(-2.7277)	(-2.7662)	(-2.4857)	(-2.6881)
Tobin's Q	0.0068***	0.0067***	0.0069***	0.0072***	0.0065***	0.0067***
	(3.8742)	(3.9373)	(4.0590)	(3.8964)	(3.7242)	(3.8819)
Log(Employees)	0.0043	0.0048*	0.0046*	0.0050*	0.0051*	0.0049*
	(1.5297)	(1.7497)	(1.6853)	(1.9000)	(1.9232)	(1.7530)
ROA(f,t−1)	0.5669***	0.5590***	0.5629***	0.5564***	0.5433***	0.5581***
	(12.3009)	(11.2891)	(11.5363)	(12.5765)	(11.0488)	(11.2384)
Constant	-0.0263	-0.0229	-0.0240	-0.0185	-0.0214	-0.0227
	(-1.1423)	(-1.0137)	(-1.0634)	(-0.8456)	(-0.9614)	(-1.0049)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Clustered by	Firm	Firm	Firm	Firm	Firm	Firm
Observations	4149	4149	4149	4149	4149	4149
R^2	0.4730	0.4712	0.4735	0.4909	0.4906	0.4710
Adj. R^2	0.4250	0.4231	0.4255	0.4445	0.4443	0.4229



Figure 1: Event Study — IHS(Commits)

The figure reports results from the baseline event-study specification in Section 4.1 in event time with 95% confidence interval. The red dotted vertical line indicates the month of the job transition (the start of the engineer's new job). On the x-axis, we report the months relative to the job transition. On the y-axis, we report *IHS(Commits)*.

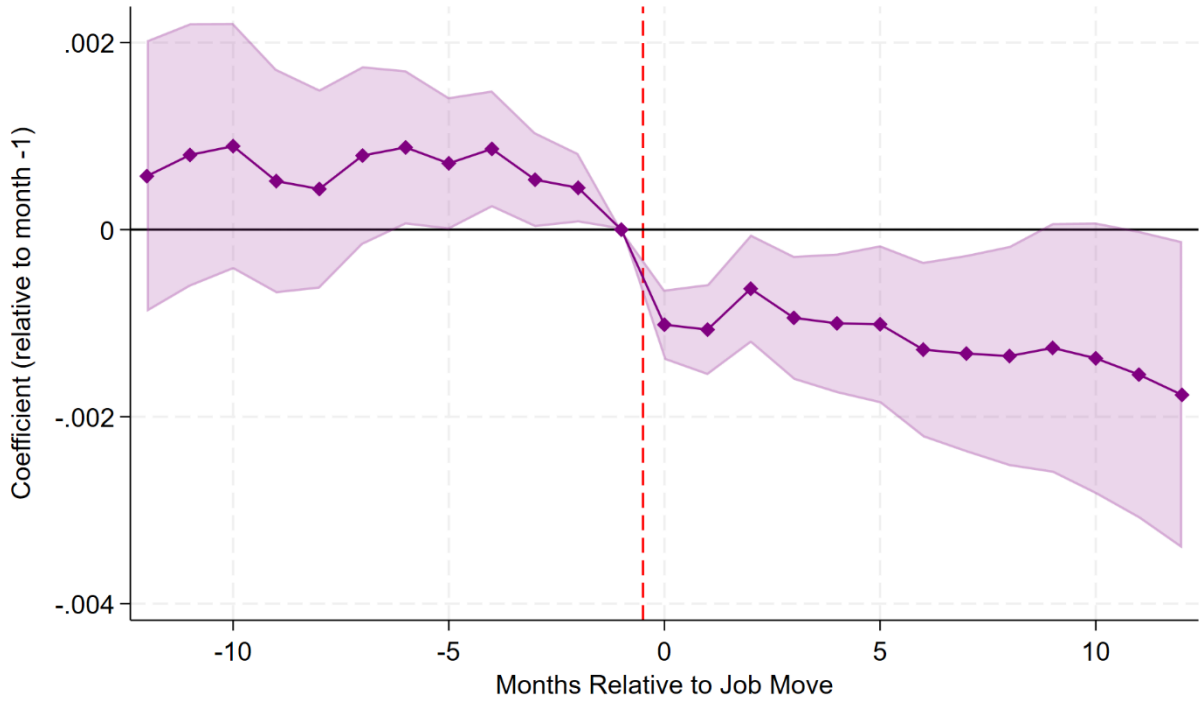


Figure 2: Event Study — IHS(PRs Merged)

The figure reports results from Eq. (1) in event time with 95% confidence interval. The red dotted vertical line indicates the month of the job transition (the start of the engineer's new job). On the x-axis, we report the months relative to the job transition. On the y-axis, we report IHS(PRs Merged).

Online Appendix

Appendix Table A1 (LPM): Mobility and Signaling Activity — Extensive Margin

This table reports linear-probability event-study estimates of the extensive margin of signaling activity around job transitions. The dependent variables are indicators equal to one if the engineer records any commit, any pull request opened, or any repository created on a given day, so coefficients are in probability points. Panel A reports grouped event-time windows relative to the far-pre window [−360, −181]; Panel B reports the pre/post specification relative to the omitted post-move window. "Base rate, far pre" is the mean active-day rate in the far-pre window, and "Pct. lift, near pre" equals 100 × the near-pre coefficient divided by that base rate. Please refer to Appendix Table A4 for detailed variable definitions. All specifications include event and calendar-day fixed effects. t-statistics based on standard errors clustered at the engineer level are reported in parentheses. ***, **, * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Grouped Event Windows

	1(Any Commit)	*1(Any PR Opened)*	*1(Any Repo Created)*
	(1)	*(2)*	*(3)*
Near pre [−180, −1]	0.0040***	0.0002	0.0017***
	(7.7617)	(1.2931)	(8.8305)
Near post [0, +180]	−0.0118***	−0.0012***	−0.0016***
	(−15.7086)	(−4.3550)	(−5.5842)
Far post [+181, +360]	−0.0128***	−0.0011***	−0.0017***
	(−13.0841)	(−3.3797)	(−4.6622)
Observations	24,936,197	24,936,197	24,936,197
N Engineers	22,154	22,154	22,154
R^2	0.140	0.089	0.035
Adj. R^2	0.138	0.087	0.033
Base rate, far pre [−360, −181]	0.0430	0.0066	0.0133
Pct. lift, near pre (%)	9.2	3.8	12.9
Event-ID FE	Yes	Yes	Yes
Calendar-date FE	Yes	Yes	Yes

Panel B: Pre/Post Specification

	1(Any Commit)	*1(Any PR Opened)*	*1(Any Repo Created)*
	(1)	*(2)*	*(3)*
Pre-move [−360, −1]	0.0166***	0.0015***	0.0037***
	(27.5333)	(7.3894)	(18.5307)
Observations	24,936,197	24,936,197	24,936,197
N Engineers	22,154	22,154	22,154
R^2	0.140	0.089	0.035
Adj. R^2	0.138	0.087	0.033
Base rate, far pre [−360, −181]	0.0430	0.0066	0.0133
Event-ID FE	Yes	Yes	Yes
Calendar-date FE	Yes	Yes	Yes

Appendix Table A2: Engineer Background — Education, Demographics, Occupation, and Seniority

This table reports background descriptive statistics for the engineers in the position-level transition sample (22,789 workers, matched by user_id to Revelio profile and education records). Panels A–C report education (highest degree, field of study, and top universities); Panels D–G report demographics (estimated age, sex, ethnicity, and LinkedIn connections); Panel H reports the occupation composition; and Panel I reports the seniority distribution. Estimated age is the earliest education start year minus 18; sex and ethnicity are Revelio model predictions; and LinkedIn connections are top-coded at 500. In Panel B, shares are computed over the engineers with a reported field of study.

Panel A: Highest degree

Highest degree	N	%
Master	8,939	43.0
Bachelor	8,791	42.3
Doctor	1,569	7.6
MBA	860	4.1
Associate	359	1.7
High School	260	1.3
Total	20,778	100.0

Panel B: Field of study (highest degree)

Field of study	N	%
Engineering	11,354	71.6
Business	2,491	15.7
Information Technology	344	2.2
Mathematics	320	2.0
Economics	296	1.9
Statistics	211	1.3
Physics	210	1.3
Finance	129	0.8
Biology	103	0.6
Education	97	0.6
Chemistry	88	0.6
Marketing	72	0.5
Accounting	50	0.3
Law	45	0.3
Total	15,810	99.7

Panel C: Top universities (by number of engineers)

University	N (engineers)
Udacity	816
Georgia Institute of Technology	687
University of California Berkeley	633
University of Washington	480
Stanford University	479

University of Southern California	440
The University of Texas at Austin	431
San Jose State University	415
Carnegie Mellon University	385
University of Mumbai	375
University of California San Diego	367
Arizona State University	367
University of Illinois at Urbana-Champaign	359
The University of Texas at Dallas	341
Northeastern University	334
New York University	327
Columbia University in the City of New York	281
Harvard University	274
University of California Los Angeles	262
Massachusetts Institute of Technology	259
Total (top 20 shown)	8,312

Panel D: Estimated age at first job move

Statistic	Estimated age at first job move
N	20485
Mean	29.3
SD	7.6
P25	24
Median	27
P75	33

Panel E: Sex (predicted)

Sex (predicted)	N	%
Male	18,351	80.5
Female	4,326	19.0
empty	112	0.5
Total	22,789	100.0

Panel F: Ethnicity (predicted)

Ethnicity (predicted)	N	%
White	11,632	51.0
API	7,696	33.8
Hispanic	1,819	8.0
Black	1,477	6.5
Multiple	125	0.5

Native	29	0.1
	11	0.0
Total	22,789	100.0

Panel G: LinkedIn connections

Statistic	LinkedIn connections
N	22783
Mean	474
P25	420
Median	500
P75	500
P90	500

Panel H: Sample composition by occupation — top 15 detailed occupations (ONET 6-digit)

ONET	Occupation	N	%
15-1252	Software Developers	20,925	38.72
15-2051	Data Scientists	5,402	10.0
17-2199	Engineers, All Other	2,119	3.92
15-1299	Computer Occupations, All Other	2,115	3.91
15-1251	Computer Programmers	1,595	2.95
15-1241	Computer Network Architects	1,583	2.93
15-1243	Database Architects	1,179	2.18
15-1253	Software QA Analysts & Testers	1,165	2.16
15-1254	Web Developers	842	1.56
15-1232	Computer User Support Specialists	659	1.22
17-2141	Mechanical Engineers	656	1.21
15-1211	Computer Systems Analysts	613	1.13
19-1011	Animal Scientists / Research	592	1.1
13-1111	Management Analysts	589	1.09
27-2012	Producers & Directors (incl. recruiters proxy)	578	1.07
	Total (top 15 shown)	40,612	75.15

Panel I: Engineer characteristics — seniority distribution (1–7)

Seniority	N	%
1	11,013	20.38
2	21,969	40.66
3	9,067	16.78
4	6,242	11.55
5	4,633	8.57
6	1,030	1.91

7	83	0.15
Total	54,037	100.0

Appendix Table A3: Firm-Level Consequences — Future Headcount Conditional on Current Headcount Changes and Employee GitHub Activity

This table reports OLS estimates relating future headcount changes to the current headcount change, employee GitHub disclosure intensity (ShareSignaling), and their interaction, at the t+1 and t+2 horizons. Columns (1)–(2) use the full sample, Columns (3)–(4) the shrinking-firm subsample (ΔHeadcount < 0), and Columns (5)–(6) the growing-firm subsample (ΔHeadcount > 0). All specifications include industry-by-year fixed effects and control for log total assets, leverage, cash ratio, Tobin's Q, lagged ROA, and revenue growth. Please refer to Appendix Table A4 for detailed variable definitions. t-statistics based on standard errors clustered at the firm level are reported in parentheses. ***, **, * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Sample	*Full Sample*		*Shrinking (ΔHC<0)*		*Growing (ΔHC>0)*	
Dep. var	*Headcount*	*Headcount*	*Headcount*	*Headcount*	*Headcount*	*Headcount*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
	t+1	*t+2*	*t+1*	*t+2*	*t+1*	*t+2*
ΔHeadcount(f,t)	0.0477	0.0921**	-0.2184	-0.0692	0.0861**	0.0843
	(0.9175)	(2.0259)	(-1.5129)	(-1.1694)	(2.1613)	(1.3191)
ShareSignaling(f,t)	-0.0289**	-0.0075	0.0144	0.0194	-0.0440***	-0.0100
	(-2.0830)	(-0.6256)	(0.4720)	(0.7957)	(-2.7049)	(-0.5183)
ΔHC × ShareSignaling	0.2231**	0.1045	0.3968	0.2743**	0.2729***	0.0930
	(2.0617)	(1.1836)	(1.3408)	(2.0510)	(3.6029)	(0.8150)
Log(total assets)	-0.0053*	-0.0052*	-0.0064	-0.0081**	-0.0013	-0.0025
	(-1.9460)	(-1.9100)	(-1.2226)	(-2.0429)	(-0.3987)	(-0.7089)
Leverage	-0.0609***	-0.0359*	-0.0610	-0.0950***	-0.0568**	0.0100
	(-2.8871)	(-1.6868)	(-1.2447)	(-3.0473)	(-2.4825)	(0.3146)
Cash / assets	0.0854**	0.0778**	-0.0778	0.0078	0.1403***	0.1205***
	(2.3472)	(2.0344)	(-1.1715)	(0.1145)	(4.2724)	(2.9320)
Tobin's Q	0.0166***	0.0140***	0.0186**	0.0109*	0.0137***	0.0129***
	(6.7131)	(5.4081)	(2.4296)	(1.9322)	(6.8784)	(4.3475)
Lagged ROA	0.0321	0.0002	0.1544	0.0087	-0.0213	-0.0043
	(1.1504)	(0.0115)	(1.3927)	(0.2303)	(-0.7616)	(-0.0804)
Revenue growth	0.0055	0.0008	0.0282	-0.0082	0.0045	0.0004
	(1.2880)	(0.5091)	(0.5629)	(-0.4438)	(1.1470)	(0.2563)
Constant	0.0530*	0.0477*	0.0303	0.0746*	0.0211	0.0178
	(1.9511)	(1.6973)	(0.5667)	(1.7787)	(0.6094)	(0.4797)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Clustered by	Firm	Firm	Firm	Firm	Firm	Firm
Observations	4084	3150	1287	856	2488	2041
R^2	0.2023	0.1864	0.2382	0.1933	0.2365	0.2071
Adj. R^2	0.1285	0.1035	0.0870	0.0033	0.1489	0.1107

Appendix Table A4: Variable Definitions and Data Sources

Variable	Definition	Source
IHS(Commits)	IHS of the number of code commits authored by the engineer on a given day. Commits capture changes or additions to code and proxy the intensity of public coding effort. Inverse hyperbolic sine (IHS) of the daily count, sinh-inverse(x)=ln(x+sqrt(x^2+1)); defined at zero, behaves like log in the upper tail. Daily engineer-level count.	GitHub Archive
IHS(PRs Opened)	IHS of the number of pull requests opened by the engineer on a given day. Pull requests opened capture attempts to contribute code to repositories (often involving external review). Inverse hyperbolic sine (IHS) of the daily count, sinh-inverse(x)=ln(x+sqrt(x^2+1)); defined at zero, behaves like log in the upper tail. Daily engineer-level count.	GitHub Archive
IHS(Repos Created)	IHS of the number of new public repositories created by the engineer on a given day. Repository creation captures deliberate public presentation of a new project or technical artifact. Inverse hyperbolic sine (IHS) of the daily count, sinh-inverse(x)=ln(x+sqrt(x^2+1)); defined at zero, behaves like log in the upper tail. Daily engineer-level count.	GitHub Archive
IHS(Forks Received)	IHS of the number of times other users forked (copied) the engineer's repositories on a given day. Captures whether disclosed code is inspected, modified, or built upon by others. Inverse hyperbolic sine (IHS) of the daily count, sinh-inverse(x)=ln(x+sqrt(x^2+1)); defined at zero, behaves like log in the upper tail. Daily engineer-level count.	GitHub Archive
IHS(Active Forks)	IHS of the number of forks of the engineer's repositories that show subsequent commit activity (i.e., forks that are actively used rather than merely copied) on a given day. Inverse hyperbolic sine (IHS) of the daily count, sinh-inverse(x)=ln(x+sqrt(x^2+1)); defined at zero, behaves like log in the upper tail. Daily engineer-level count.	GitHub Archive
IHS(PRs Merged)	IHS of the number of pull requests merged into the engineer's repositories on a given day, capturing outside contributions accepted into the engineer's code (collaboration / integration). Inverse hyperbolic sine (IHS) of the daily count, sinh-inverse(x)=ln(x+sqrt(x^2+1)); defined at zero, behaves like log in the upper tail. Daily engineer-level count.	GitHub Archive
Δ IHS(activity), post-pre	Cross-sectional 2SLS outcome: change between post- and pre-period averages of the IHS-transformed daily activity (average over [0,+12] months minus average over [-12,-1] months).	GitHub Archive
Event time	Days relative to the move date, defined as the start date of the engineer's new job. Engineer-by-day panel spans [-360, +360] around each transition; transitions near the start/end of the GitHub sample have shorter observed windows.	Revelio Labs
Far pre [-360, -181]	Omitted baseline window in the grouped event study: engineer-days 181-360 days before the move date.	Revelio Labs
Near pre [-180, -1]	Indicator for engineer-days in the 180 days immediately before the move date (1 if the day falls in [-180, -1], else 0).	Revelio Labs
Near post [0, +180]	Indicator for engineer-days in the 180 days from the move date onward.	Revelio Labs
Far post [+181, +360]	Indicator for engineer-days 181-360 days after the move date.	Revelio Labs
Pre-move [-360, 0]	Indicator equal to one for engineer-days in the pre-move window (up to 360 days before through the move date) in the pre/post specification.	Revelio Labs
PreMove	Indicator equal to one for engineer-days in the pre-move window relative to the transition; interacted with worker/firm moderators in the heterogeneity tests.	Revelio Labs
Leaver	Endogenous treatment: indicator equal to one if the engineer departs the origin firm following the contraction period. Instrumented by firm-level headcount contraction.	Revelio Labs
Seniority (1-7)	Revelio seniority level on a 1-7 scale (career stage; higher = more senior).	Revelio Labs
Junior	Indicator equal to one if the engineer's origin-firm seniority is at or below 3 on the Revelio 1-7 scale (i.e., below Manager); interacted with PreMove in the heterogeneity tests.	Revelio Labs

Total compensation	Estimated total annual compensation ($000s).	Revelio Labs
High compensation	Indicator equal to one if the engineer's total compensation is above the sample median; interacted with PreMove in the heterogeneity tests.	Revelio Labs
COVID	Indicator equal to one for engineer-days falling in the COVID-19 period (March 2020-December 2021); interacted with PreMove in the heterogeneity tests.	Revelio Labs
ShareSignaling	Share of a firm's engineering workforce with above-median public GitHub activity in a given year	Revelio Labs
log(firm size)	First-stage control for origin-firm size (log employment).	Compustat
ΔEarnings(f,t)	Contemporaneous change in firm earnings in year t; key regressor in the earnings-persistence specification.	Compustat
ROA	Return on assets (firm-year). Dependent variable / control in the firm-level earnings analyses.	Compustat
Op.ROA	Operating return on assets (firm-year); operating-performance dependent variable.	Compustat
ROA(f,t-1)	Lagged (prior-year) return on assets; control for persistence in returns.	Compustat
Log(Total Assets)	Natural log of firm total assets; firm-level control.	Compustat
Leverage	Firm leverage, debt divided by total assets.	Compustat
Cash/Assets	Cash ratio: cash divided by total assets.	Compustat
Tobin's Q	Tobin's Q (market value relative to book value of assets); firm-level control.	Compustat
Employees	Origin-firm employees (000s).	Compustat
Log(Employees)	Natural log of the number of firm employees; firm-level control.	Compustat
Headcount contraction (continuous)	Continuous firm-level instrument measuring the year-over-year employment decline at the origin firm (from Compustat emp). Higher values denote larger contractions.	Compustat
Headcount decline >= X%	Binary instruments equal to one if the firm's annual headcount declined by at least the threshold X% year-over-year (X in {10, 15, 20, 25, 30}); -25% is the preferred threshold. Computed from Compustat emp.	Compustat
Total assets	Origin-firm total assets ($M).	Compustat
R&D / Sales	Origin-firm research-and-development expense divided by sales (R&D intensity).	Compustat
High R&D intensity	Indicator equal to one if the origin firm's R&D-to-sales ratio is above the sample median; interacted with PreMove in the heterogeneity tests.	Compustat