

## Research Question

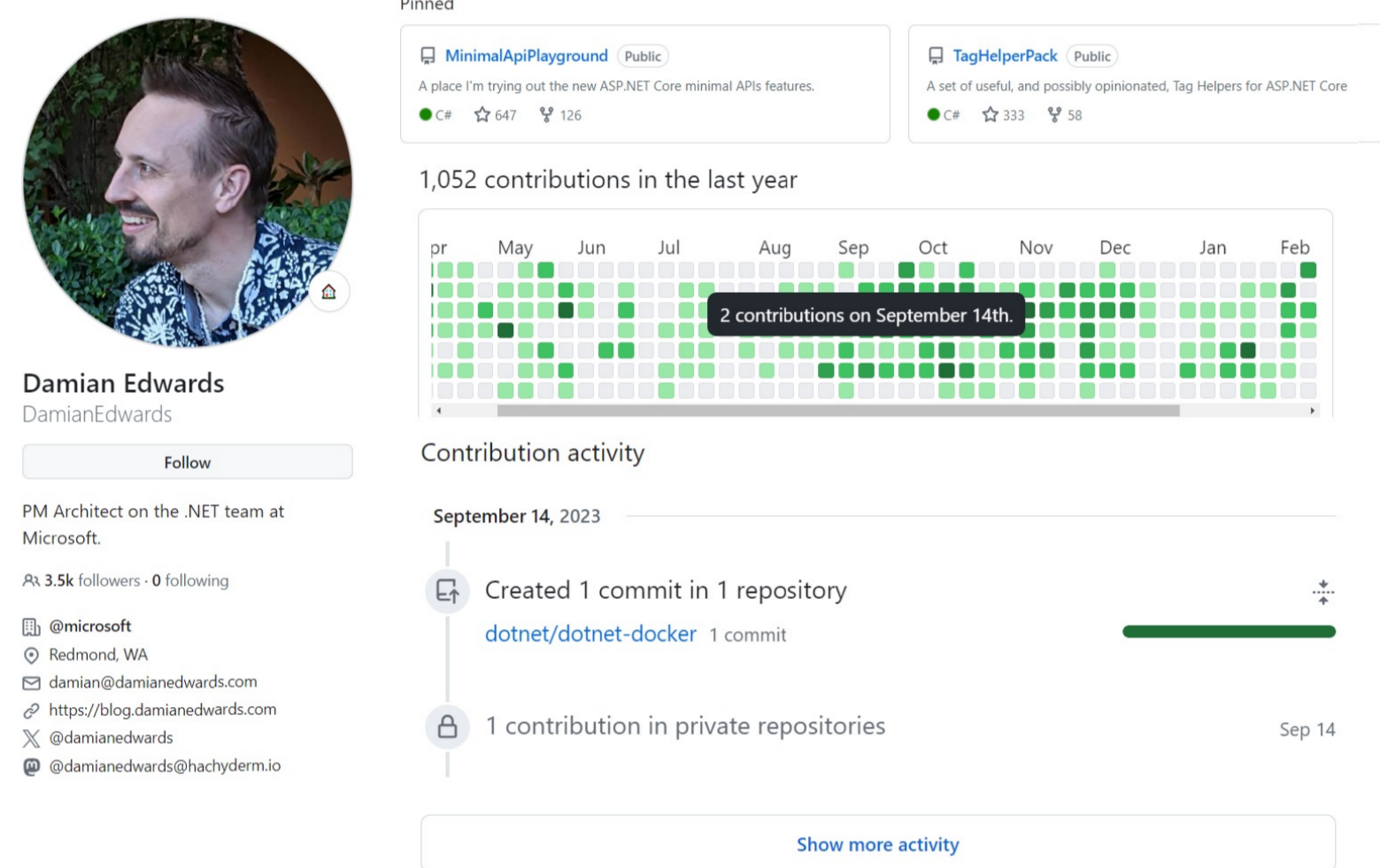
Over **80%** job-seekers and **90%** employers use web platforms for job search and recruitment. Some of these platforms offer an abundance of data on **real-time employee output**, reducing information asymmetry on employee quality.

Through the new information on employee quality, how does big data on web platforms influence labor reallocation, and subsequently, what are the effects on firms?

## Setting: GitHub–Hub of Source Code & Technical Talent

GitHub is a software management and version control platform, used by over 100 million world-wide software developers and 200,000 firms in the United States alone.

- GitHub stores all data and codes in online **repositories**, providing users the flexibility to choose the visibility of their repositories – whether **public**, allowing anyone to view and download the code, or **private**, restricting access solely to selected users (for instance, within the same firm)

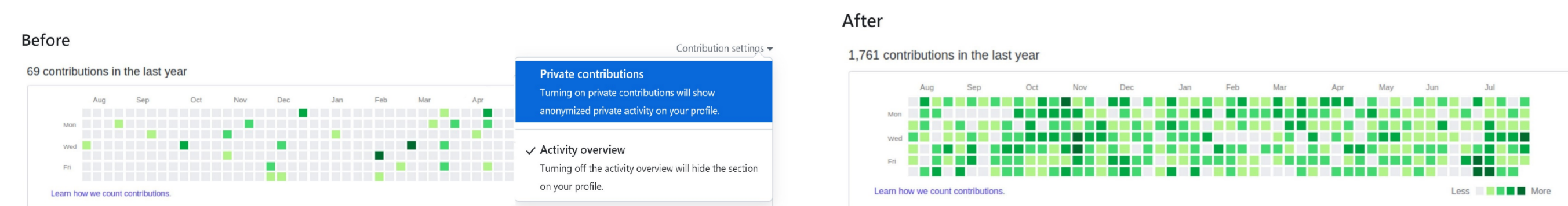


- GitHub publicly displays **user's contributions** (in the form of modifications) to GitHub repositories, providing their real-time high-frequency work output
- GitHub contributions are **correlated with career advancements** and publications for individuals and **organizational capital** and startup financing for firms
- Recruiters increasingly rely on GitHub user activity to source and evaluate potential tech-sector employees

## Identification Strategy: GitHub Allows Sharing Private Contributions

Before **May 19, 2016**, users could only display contributions made to public GitHub repositories on their profiles. However, on May 19, 2016, GitHub introduced the option for users to **include the number of (anonymized) private contributions** to their profiles.

- Users could instantaneously show their **historical** private contributions with a click
- We use users' pre-policy change private contributions to construct an exogenously available measure of their output



## Data & Sample Construction

We map US-based GitHub users, with a public GitHub and LinkedIn profile, to their employment records from LinkedIn. Our sample covers over **380,000 tech workers** employed in the US and active on GitHub before May 2016. We download users' monthly contributions from GitHub API and GitHub Archive.

## Empirics: Measuring Employee Productivity using the AKM

To disentangle employee productivity from firm productivity, we exploit employees' moves across firms to isolate individual productivity (**employee fixed effects**) using AKM 1999 [1]:

$$Y_{i,f,t} = \vartheta_i + \vartheta_f + \vartheta_t + \vartheta_k + X_{i,f,t} + \varepsilon_{i,f,t}$$

- $Y$  is observed employee  $i$ 's output in role  $k$  at firm  $f$  at time  $t$
- $X_{i,f,t}$  represents time-varying experience at current firm and overall work experience
- $\vartheta_f$  is the unobserved firm effect;  $\vartheta_t$  is time fixed effect
- $\vartheta_i$  is the unobserved individual effect, our measure of employee productivity

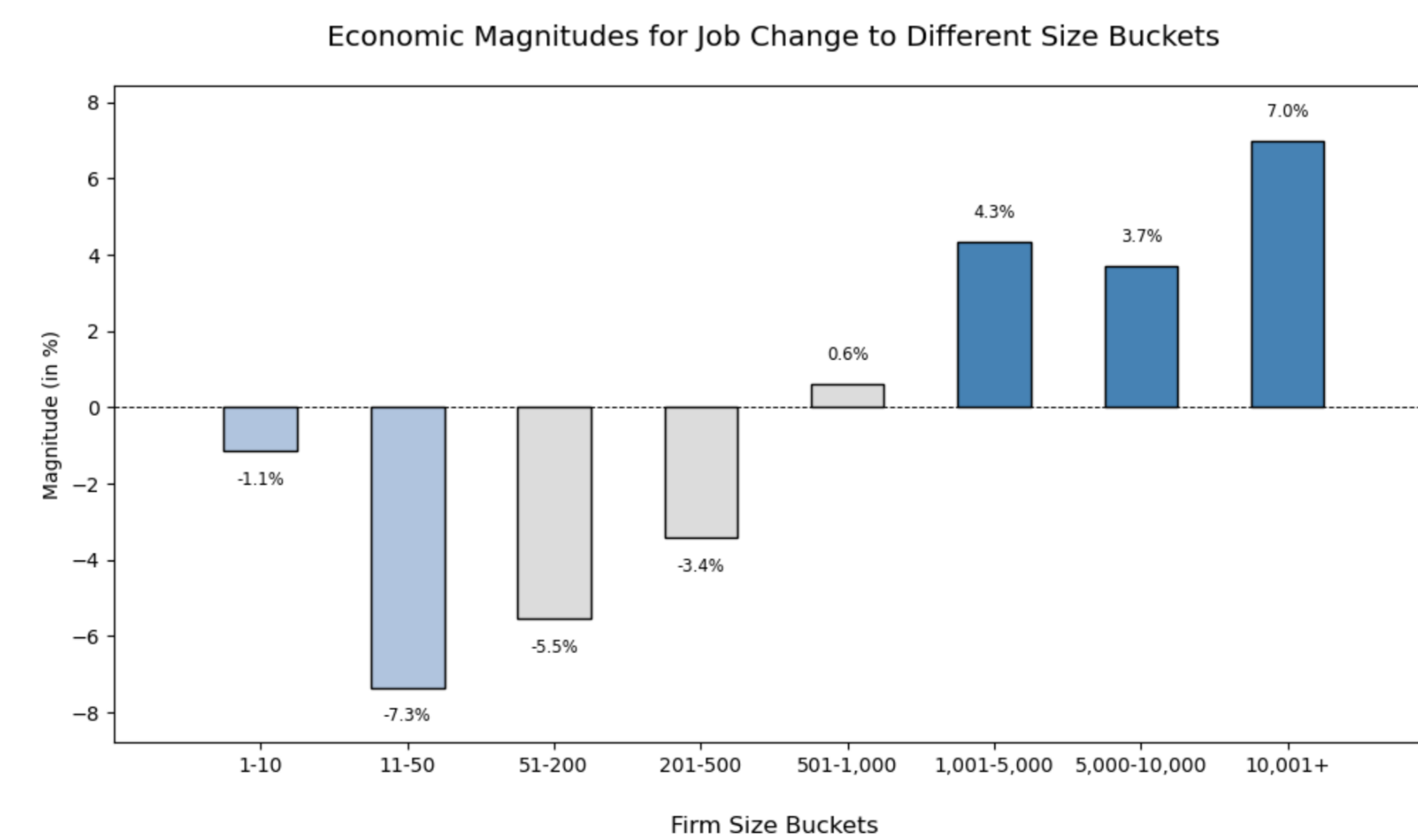
We use monthly panel data on employees' private contributions and their employment details in the **5 years prior to the shock** (May 2011-April 2016) in the AKM estimation.

## Main Result: Talent Reallocation towards Large Firms

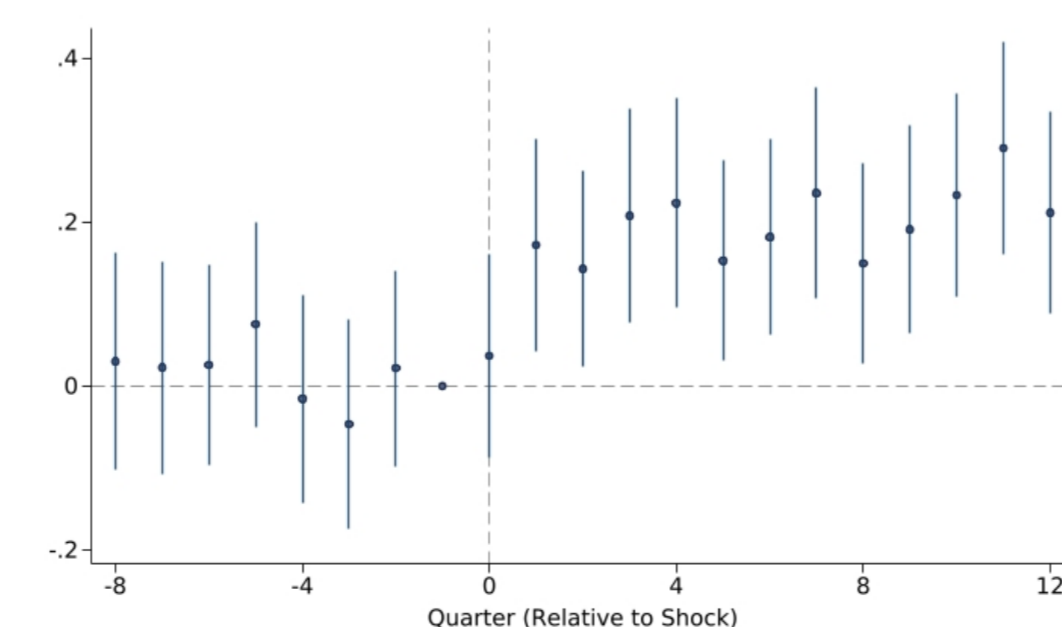
With an employee-firm panel three years before and after GitHub's policy change, we carry out a DID analysis where we compare the change in employee mobility for individuals with higher AKM-implied productivity after the GitHub policy change, as compared to those with lower productivity:

$$1(\text{Job Change}_{i,t}) = \beta \text{Productivity}_i \times 1(\text{Post}_t) + \lambda X_{i,t} + \alpha_i + \delta_t + \gamma_{i,t} + \varepsilon_{i,t}$$

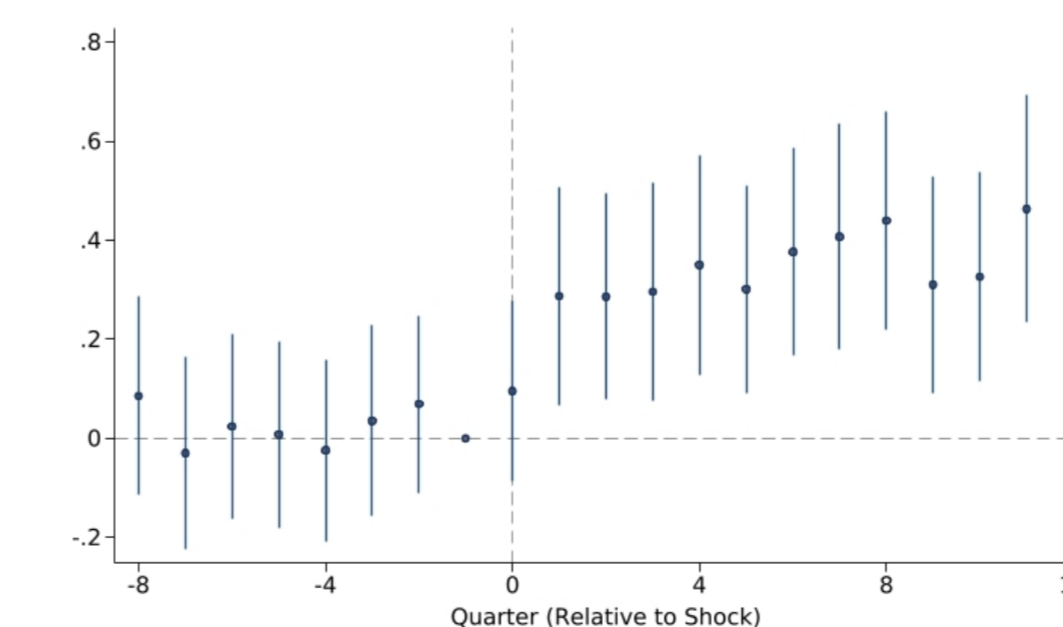
- $1(\text{Job Change}_{i,t})$  is 1 if employee  $i$  moves to a new firm in quarter  $t$  and 0 otherwise
- $\text{Productivity}_i$  is employee  $i$ 's normalized pre-treatment AKM-derived productivity estimate
- $\text{Post}_t$  is 1 in the 12 quarters after the shock and 0 in the 12 quarters before
- $X_{i,t}$  captures employee  $i$ 's time-varying seniority and salary;  $\alpha_i$  is employee fixed effects;  $\delta_t$  is time fixed effects
- $\gamma_{i,t}$  includes time-varying cohort-language-role fixed effects as well as employee's industry-location-time fixed effects



Productive Employees Move Less to Small Firms and More to Large Firms



Productive Employees Move to Large Firms (> 1,000 Employees)



Productive Employees Move From Small Firms (<= 50 Employees) to Large Firms (> 1,000 Employees)

## Additional Results

### Employees benefit from moving to large firms

- Moves to large firms are associated with increase in salary and/or seniority
- Large firms retain productive employees through internal promotions

### GitHub policy change democratizes screening of talent

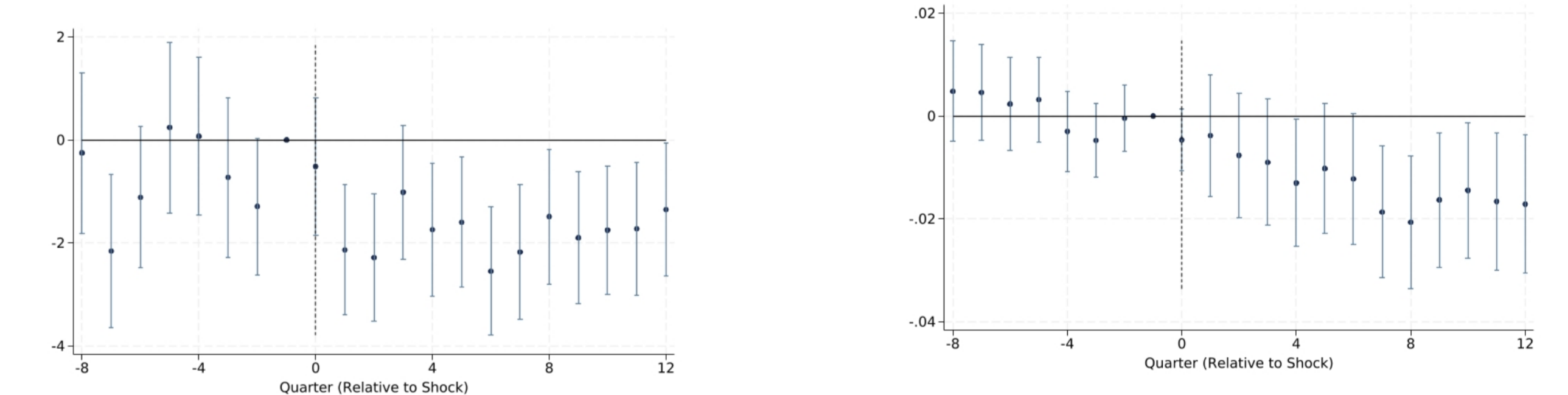
- Stronger response for individuals with below median work experience at the time of the change, below median schooling, and who graduate from non-elite universities

### Large firms gain both due to employee-preference and firm-demand

- Productive employees more likely to move to firms with tech-savvy HR department, more GitHub presence, and those offering steeper career growth
- Overall, productive employees match to higher growth large firms

## Firm-level Results: Growth & Productivity Drop at Impacted Small Firms

We collapse employee productivity at the firm level: **Impacted** are firms with above median average employee productivity at the time of shock. We then test for firm employment growth and productivity at impacted **small firms** in a triple difference estimation.



Employment Growth Declines at Impacted Small Firms

Prob[Public Contributions in Top Decile] Declines at Impacted Small Firms

At a more aggregate level, we find that more **impacted industries** become **6-14% more concentrated** in the five years following GitHub's policy change.

## Conclusion

- Web platforms reduce information asymmetries in the labor market, which benefits productive workers, hurts small firms, and benefits large firms
- Web platforms have contributed to increased labor concentration at large firms

## References

[1] John M Abowd, Francis Kramarz, and David N Margolis. High wage workers and high wage firms. *Econometrica*, 67(2):251–333, 1999.

## Interested in More Details?



- Scan the QR code to read the latest draft of our paper
- Attend our paper presentation and discussion at the AFA: Saturday, Jan. 4, 2:30 PM - 4:30 PM at San Francisco Marriott Marquis, Yerba Buena Salon 1 & 2 (Session G3: Real Effects of Information and Regulatory Frictions)