

Return to Office and the Tenure Distribution

David Van Dijcke^{1,2}, Florian Gunsilius¹, and Austin Wright³

¹Department of Economics, University of Michigan

²Risk Analytics Division, Ipsos Public Affairs

³Harris School of Public Policy, University of Chicago

May 7, 2024

Abstract

With the official end of the COVID-19 pandemic, debates about the return to office have taken center stage among companies and employees. Despite their ubiquity, the economic implications of return to office policies are not fully understood. Using 260 million resumes matched to company data, we analyze the causal effects of such policies on employees' tenure and seniority levels at three of the largest US tech companies: Microsoft, SpaceX, and Apple. Our estimation procedure is nonparametric and captures the full heterogeneity of tenure and seniority of employees in a distributional synthetic controls framework. We estimate a reduction in counterfactual tenure that increases for employees with longer tenure. Similarly, we document a leftward shift in the seniority distribution towards positions below the senior level. These shifts appear to be driven by employees leaving to larger firms that are direct competitors. Our results suggest that return to office policies can lead to an outflow of senior employees, posing a potential threat to the productivity, innovation, and competitiveness of the wider firm.

JEL Codes: J21, J62, M54, C14

Keywords: causal inference, distributional synthetic controls, human capital, return to office, seniority, tenure, work from home

{dvdijcke;ffg}@umich.edu, austinlw@uchicago.edu. We thank Mel Stephens and Germano Wallerstein for helpful discussion and comments as well as Ewan Rawcliffe for outstanding research assistance. The authors gratefully acknowledge support from the Peterson G. Peterson Foundation Pandemic Response Policy Research Fund and the Becker Friedman Institute. FG gratefully acknowledges support from a MITRE research award from the University of Michigan.

1 Introduction

Since the World Health Organization officially declared an end to the COVID-19 health crisis, a main focus of companies has shifted towards creating workplace policies for a post-pandemic economy. In this, the preferences of employers and employees often seem diametrically opposed. That opposition has expressed itself most notably in the debate around the effects of the return to office (RTO). This debate is far from settled: as of June 2023, an estimated 75% of tech and 28% of all companies in the United States were still “fully flexible” – either fully remote or with a voluntary in-office option (Scoop, 2023). Most business executives expect this number to grow (Bloom, Barrero, Davis, Meyer and Mihaylov, 2023a). In this paper, we provide causal evidence that RTO mandates at three large tech companies—Microsoft, SpaceX, and Apple—had a negative effect on the tenure and seniority of their respective workforce. In particular, we find the strongest negative effects at the top of the respective distributions, implying a more pronounced exodus of relatively senior personnel.

Our ability to account for the heterogeneity in the organizational structure of companies is of fundamental importance in this setting. To do so, we estimate the composition of the workforce in the counterfactual reality where the company did not implement an RTO mandate. We document changes in the employee composition that can profoundly impact a company, especially since these changes occur at the top of the hierarchy. Senior employees and those that have been at a company for a long time possess invaluable human capital and tend to have elevated productivity levels (Lazear et al., 2015), which they take with them when leaving the company. They also represent a significant investment in terms of hiring and training costs (Blatter et al., 2012). Indeed, the central tenet of resource-based theory, one of the most prevalent and empirically well-supported human capital theories, is that human capital is key to explaining why some firms outperform others (Crook et al., 2011). Changes in the employee distributions can, moreover, have ripple effects in the larger economy, by affecting the distribution of skilled workers across different types of companies (Lise and Robin, 2017; Eeckhout and Kircher, 2018).

We uncover these results by estimating the distributional causal effects of a return to office mandate on the makeup of a company’s workforce using a synthetic controls approach (Abadie and Gardeazabal, 2003; Abadie et al., 2010). Since our object of interest is the heterogeneity in the composition of the workforce of a company, we employ the distributional synthetic controls estimator proposed in Gunsilius (2023). This estimator is ideally suited for this setting, since it does not require observing the same individuals over time and as such allows for the changes in workforce composition that we are interested in. Furthermore, it provides a nonparametric estimate of the effect of an RTO on the entire tenure and seniority distribution of a target company. As a theoretical contribution, we propose a modified distributional permutation test and provide conditions for the validity of bootstrap-based uniform confidence bands.

Our case-study approach focuses on Microsoft, Apple, and SpaceX because they were among the first large American tech companies to implement RTO mandates. In particular, they did so before a wave of layoffs started hitting the tech industry from late 2022 onwards, allowing us to cleanly disentangle the causal effects of the RTO mandates. Our synthetic controls approach, moreover, allows us to study in detail what happens *within* each company after an RTO, which would not be possible with an aggregate cross-company approach. This

guarantees the internal validity of our estimates. Their external validity is twofold. First, we consider some of the largest companies in a sector where the discourse over the return to office was most heated. Together, the three companies account for over 2% of employment in the tech sector and 30% of its revenue. What happens at these companies matters for the American economy, and sets the precedent for the wider debate around the return to office. Second, we estimate nearly identical effects for all three companies despite their markedly different corporate culture and product gamut, suggesting the effects are driven by common underlying dynamics.

Our paper contributes to several branches of the literature. First, it addresses the relatively understudied impact of return-to-office (RTO) mandates. For instance, [Ding and Ma \(2023\)](#) analyze the determinants and effects of RTO mandates for a sample of S&P 500 companies by examining company balance sheets and Glassdoor reviews. They suggest that RTOs are used by managers to reassert control, which in turn reduces employee satisfaction. Using a voluntary survey of managers and their employees, [van Triest \(2023\)](#) demonstrates that work-from-home (WFH) levels are 50% lower when managers have the discretion to set RTO policies, as opposed to being constrained by organization-wide rules. [Barrero et al. \(2021\)](#) report that four out of ten employees working from home would find a new job under an RTO mandate, based on a large-scale survey of American workers. Several business and industry reports have also studied RTO. For example, a large-scale survey by [Unispace \(2023\)](#) found that almost half of firms with mandated RTOs experienced higher than usual difficulties in retaining employees and lost “key” staff members, in line with our findings. Our paper complements and extends this literature by estimating a significant outflow of more senior employees after the implementation of an RTO mandate. Unlike these previous studies, we rely on large-scale resume data rather than surveys, allowing us to study the realized consequences of RTO policies rather than self-reported behavior or preferences. Moreover, the richness of our data and estimation approach allows us to speak to causal rather than merely descriptive effects. It also allows us to study what happens to the organizational structure of companies that return to the office, which is a first-order effect of such a return.

Second, our paper relates to the large literature on WFH. Focusing on the impact of WFH on employee tenure and seniority, this body of work provides several insights relevant to ours.¹ [Mas and Pallais \(2017\)](#) estimate, using an experimental design, that workers had limited willingness to pay for hybrid work arrangements before the pandemic, with a fat tail of high-value workers, which are generally more senior. [Emanuel et al. \(2023\)](#) find that proximity to coworkers in the office increases the feedback software engineers receive but reduces their programming output, especially for senior engineers. [Hansen et al. \(2023\)](#) study hybrid work options in large-scale job postings data and find substantial heterogeneity within occupation and even company. [Mischke et al. \(2023\)](#) and [Barrero et al. \(2023b\)](#) find that office attendance is lower in larger firms, in line with our finding that employees leave for such firms at higher rates after an RTO. [Barrero et al. \(2022\)](#) find that a shift to remote work puts downward pressure on wages due to lower employee turnover, which may have reduced the college wage premium as highly educated workers are more likely to WFH ([Barrero et al.](#),

¹For a comprehensive survey of the effects of WFH, we refer to [Ding and Ma \(2023, §2.3\)](#) and [Barrero et al. \(2023a\)](#).

2023a). Finally, Bloom, Han and Liang (2023) document lower attrition rates under a hybrid work model compared to a fully-in-person one.

An important distinction between our focus on RTO mandates and the WFH literature is that we consider the effects of a *mandatory shift* to (partial) office attendance from an optional hybrid model, while the WFH literature has mostly focused on static comparisons between fully-remote, hybrid, and fully-in-person work. Moreover, while several of the papers mentioned have documented a positive relation between employee attrition rates and degree of required office attendance, our results illuminate how this attrition realizes across the employee distribution, with senior and long-tenured employees leaving at higher rates. Considering the pivotal role such employees play in a company’s development and competitive edge, along with the significant costs involved in replacing the firm-specific human capital accumulated over their tenure (Gerhart and Feng, 2021), our findings indicate that such “uneven” attrition should be a major concern for any company considering a return to office.

The rest of this paper is structured as follows. In Section 2 we discuss our empirical approach in terms of data and estimation and present some descriptive evidence; in Section 3 we present our causal results and discuss their implications in light of the academic literature; before concluding in Section 4.

2 Empirical Approach

2.1 Data

People Data Labs We study the impact of RTO across the employee distribution using data provided by People Data Labs (PDL) (<https://www.peopledatalabs.com/>) via the Dewey Data platform (<https://deweydata.io/>). The dataset consists of global resume data for nearly 1 billion unique individuals, with 1.2 billion person records in North America from over 300 million resumes. The data contains information on the start and end date of a person’s position at a given company, as well as the role (trainee, entry, partner, etc) of the position.

We estimate the tenure distribution within a company during a given quarter by calculating

$$(1) \quad \text{Tenure}_{ijt} := \begin{cases} t - T_{0,ij} & \text{if } T_{ij} > t \\ T_{ij} - T_{0,ij} & \text{else} \end{cases}$$

where t indexes the end of the quarter, $T_{0,ij}$ is the start date of employee i at company j , and T_{ij} is the end date. Then, we further sum this variable across multiple subsequent employment episodes at the same company for the same individual, to calculate total tenure for a single employment spell. This is needed because some employees list a single employment spell at the same company under separate resume items to indicate promotions.

As another proxy for seniority within the company, we use the discrete distribution of job titles in the PDL data. The lowest level is “unpaid” and the highest “CXO”, corresponding to a C-suite title. The full list of titles and their corresponding numeric codes is shown in

Table 1: Seniority Titles

Title	Unpaid	Training	Entry	Manager	Senior	Owner	Partner	Director	VP	CXO
Ranking	1	2	3	4	5	6	7	8	9	10

Note: table displays seniority levels and corresponding numeric code of employees’ positions as classified in the People Data Labs data.

Table 1. Based on this, we define an employee’s title within a given quarter as the highest title it held within the company in that quarter.

Return to Office We collect data on the return to office dates of tech firms from two sources. First, we manually collect the month a return-to-office mandate was implemented for 55 large tech firms (e.g. Google, Apple, IBM) based on news reports and leaks on a popular anonymous employee forum, Blind. Of those, 19 companies implemented an RTO mandate, while the other 36 continued to allow remote work to varying degrees as of February 2023. We depict the chronology of the return to office dates we collected in Figure 1. Though this is not necessarily a representative sample, we can see that RTO implementations only picked up more rapidly from spring 2023 onwards.

Second, we supplement this list with data scraped from the Flex Index by Scoop Technologies, a hybrid working platform (<https://www.flex.scoopforwork.com/>). These data provide information on the remote work arrangements of more than 8,000 companies. While they do not contain information on the date these companies returned to work (if at all), we leverage these data to expand our set of “never-treated” companies that never returned to the office. These serve as our control units, together with companies that were “not-yet-treated” until 6 months after the end of our sample period. We were able to match these data directly to the PDL data based on the company name and website domain, as both fields had accurate coverage in the two datasets.

Tech Layoffs At the time of writing this article, the tech sector in the United States has experienced persistent rounds of large-scale layoffs since late 2022. As these layoffs were in general independent of companies’ RTO mandates, they are likely to confound our estimates. Of course, an employee’s refusal to comply with an RTO mandate could affect their likelihood of being laid off, but this does not affect the independence of the *implementation* of both policies. To address this concern, we leverage data from layoffs.fyi, a website that has tracked the near-universe of tech sector layoffs since the start of the COVID-19 pandemic. The website aggregates information on layoffs from news reports, company announcements, and crowdsourced reports from employees themselves. We match these data to the PDL data with a combination of fuzzy string matching and manual checking to ascertain we capture all layoffs in the data.

Sample Construction We begin by identifying the relevant set of “treated units” used in the main estimation. To avoid concurrent shocks during the estimation window, we restrict the set of firms to companies without layoffs during the two quarters before and one quarter after their RTO mandate. This restriction yields three firms: Microsoft, SpaceX, and Apple.

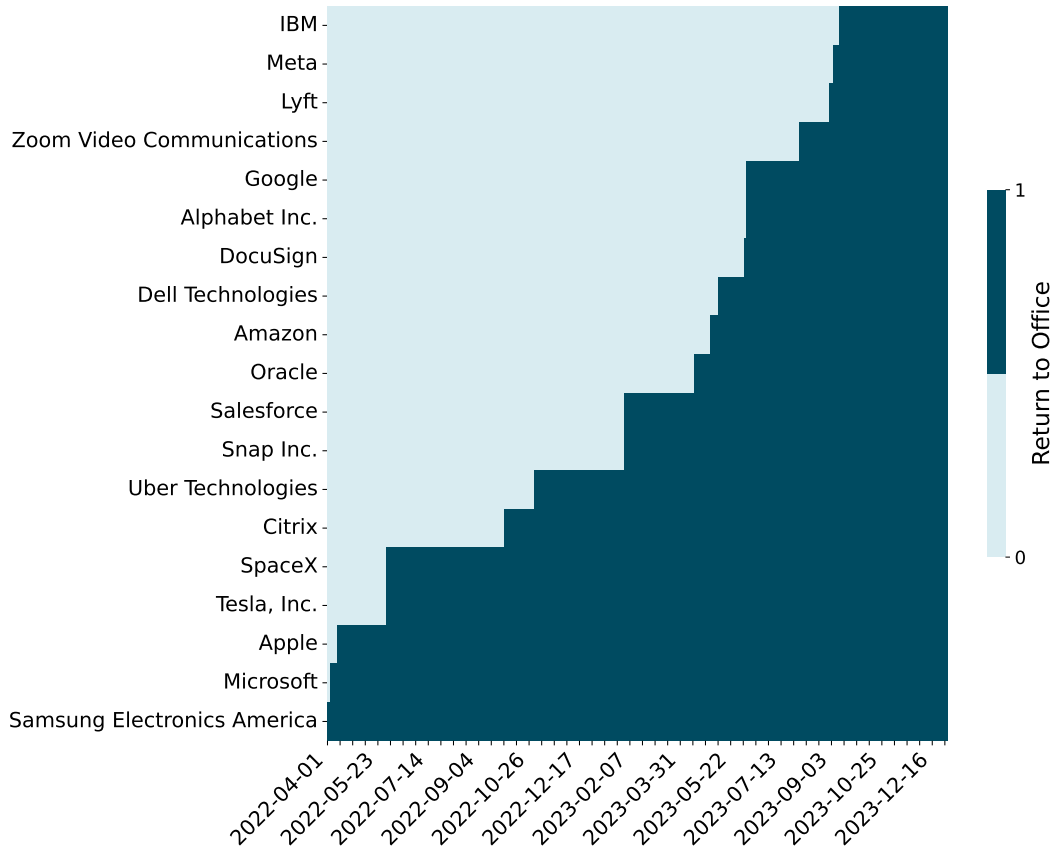


Figure 1: Return to Office Dates of Large Tech Firms

Note: figure shows return to office dates of major tech firms, ordered chronologically at a frequency of 7 weeks. Dark blue color indicates the firm has implemented an RTO mandate.

All three companies implemented their RTO mandate relatively early. They all implemented several rounds of layoffs as well, but, as mentioned, only did so at least one quarter after their return to office. Moreover, they are all large tech firms with several thousands of employees that produce both hardware and software. All three companies also shifted from an optional hybrid model to mandatory office attendance. The degree of office presence mandated varied, with Microsoft mandating 50% attendance (O’Loughlin, 2022), Apple requiring 1 day a week (O’Loughlin, 2023), and SpaceX demanding full-time attendance (40 hours per week) (Mac, 2022). We replicate our main results for these three companies, but focus on Microsoft throughout the paper as it has the largest numbers of resumes in our data, it was the first to implement an RTO among the three companies, and its RTO was most clearly announced and strict in its implementation.

For each of the three companies, we construct a sample of control units by retaining only those firms that 1) had no layoffs in the sample period of three quarters around the RTO; 2) either never implemented an RTO or implemented it at least two quarters after the end of the sample period; 3) fell under the three NAICS codes most commonly associated with

the tech industry – Information (51), Professional, Scientific and Technical Services (54), or Manufacturing (33) –; and 4) had at least 5% of the total number of resumes observed for the treated unit, to guarantee comparability. Together, these restrictions leave us with a sample size of 1,141,518 employment-by-quarter observations for 32 total companies in our sample for Microsoft. An “employment” here means a single employment of an individual at a given company, where an employee changing job can have two employments within a quarter.

2.2 Descriptives

Table 2: Summary Statistics: Variation Across Companies

Statistic	Mean	Min	Pctl(25)	Median	Pctl(75)	Max
Median Tenure	918.078	304.000	602.375	853.000	1,226.500	2,009.000
Median Title	4.844	4	5	5	5	5
Number of Resumes	12,411.410	2,764	3,468.5	4,738.5	10,707	79,570
Quarterly Turnover	0.126	0.066	0.093	0.111	0.154	0.250
Share Female	0.315	0.156	0.274	0.310	0.365	0.516
Share Leaving to Startups	0.057	0.000	0.040	0.048	0.070	0.185

Note: table shows summary statistics for main outcome variables in sample of companies comprising Microsoft and 31 control companies in quarter before Microsoft’s RTO. Summary statistics were first aggregated within companies and then the spread of their distribution across companies were reported.

The data reveal several descriptive patterns that speak to our main results. In Table 2, we report cross-company summary statistics for our main variables of interest in the quarter before Microsoft’s RTO. Specifically, we first aggregate several variables across employees within each company in our main sample and then report the dispersion across companies. This shows that the average tenure at the companies in our sample is around 2.5 years, and the title held by the median employee at most of the companies is at either the manager (4) or senior (5) level. We observe between 2,700 and 80,000 resumes at any of the firms, giving us ample statistical power to estimate the empirical tenure distribution. Moreover, the quarterly turnover is in line with estimates for the information and professional sectors from the Bureau of Labor Statistics’s Job Openings and Turnover Survey (<https://www.bls.gov/jlt/>); while the share of female employees aligns with other estimates of around 30% (Bureau of Labor Statistics and NCWIT, 2021). Finally, 5% of employees that left their company within the quarter of interest took up positions at startups, which we classify as firms with less than 50 employees.

A preview of the main descriptive results is shown in Figure 2. There, we display the observed distributions of our main outcome variables, seniority and tenure, in the quarter before (“Pre”) and after (“Post”) Microsoft’s RTO. Looking at the tenure distribution in the top row, we can see that both the probability density function and the quantile function shift leftward after the RTO, suggesting a decrease in the total days tenured across the

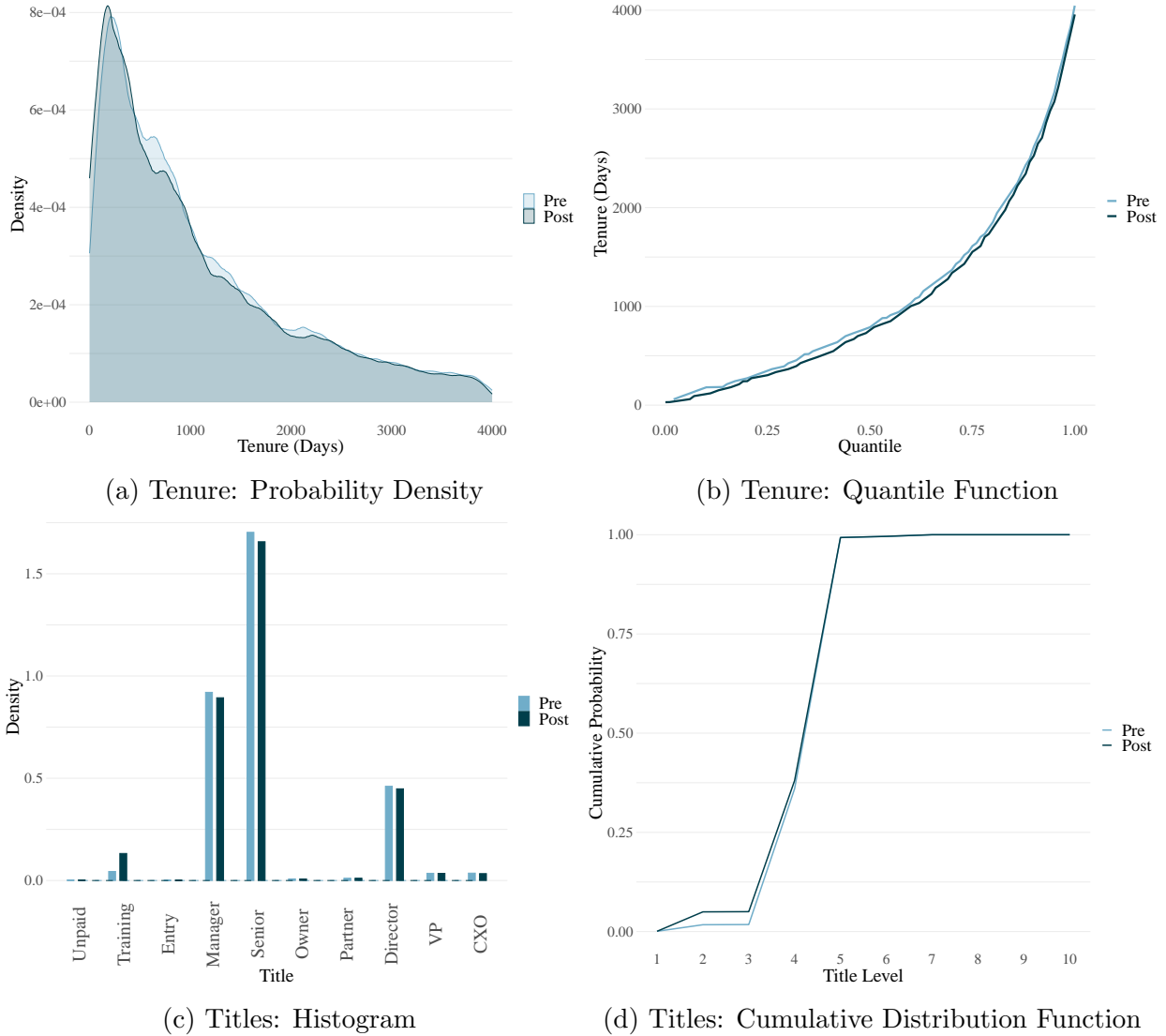


Figure 2: Distribution of Tenure and Titles

Note: 2a shows probability density functions (pdf) of tenure (in days) of Microsoft’s employees in quarter before (“Pre”, light blue) and after (“Post”, dark blue) RTO, computed using Epanechnikov kernel with bandwidth of 90 days; 2b shows empirical quantile functions of same outcome. 2c shows histogram of seniority titles; and 2d shows cumulative distribution functions of the same, with title levels mapped to ordinal variable between 1–10 in same order. Tenure distribution is restricted to bottom 90 quantiles.

distribution. Similarly, in the bottom row, we observe an increase in the histogram density for employees at the trainee level and a decrease at the senior level. This corresponds to a leftward shift in the cumulative density at seniority levels below 4 in 2d, which corresponds to the “manager” level. Overall, the patterns in the raw data already suggest a decrease in tenure and seniority at Microsoft after its RTO. Of course, these patterns are only correlational insofar as they do not account for how the distributions would have evolved in the absence of a return to office and thus fail to account for confounding shocks unrelated to the

RTO mandate. This is the purpose of our synthetic controls approach, which we describe and estimate in the next section.

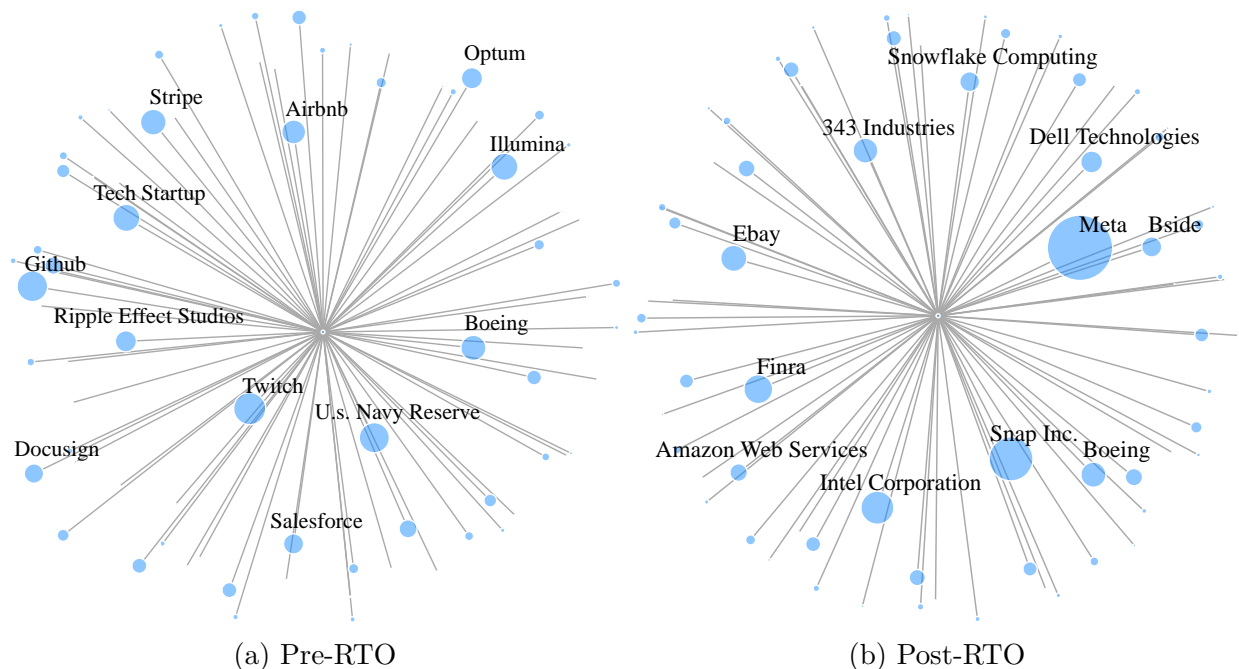


Figure 3: Quarterly Employee Outflow Network from Microsoft, by Median Tenure

Note: figure depicts flow of employees leaving Microsoft one quarter before and after their RTO. One vertex corresponds to one destination company, with the size of the vertex scaled according to the median tenure of the employees leaving for that company. Destination company names for vertices in top 20% of vertex sizes indicated by labels.

Finally, we depict the network of companies that employees leaving Microsoft moved to in Figure 3, for the quarter before and after its RTO. Each vertex indicates a company to which an employee moved. The vertex sizes correspond to the median tenure the newly hired employees had accumulated at Microsoft before leaving. Thus, the networks draw a picture of the outflow of seniority from Microsoft to its competitors. That way, one can observe a marked increase in the median tenure of employees hired at several large tech companies that compete directly with Microsoft for labor, such as Meta, Snap Inc., and Intel. In other words, after Microsoft’s RTO, several senior employees appear to have moved to large, direct competitors that did not have an RTO mandate in place. We find causal support for these descriptive patterns by estimating a counterfactual decrease in employees’ tenure and seniority at Microsoft, as well as an increase in the share of employees departing for larger companies instead of startups.

2.3 Estimation

The descriptive patterns documented above can only provide correlational evidence, as they do not account for counterfactual changes that would have affected Microsoft’s employee

distribution in the absence of an RTO. In particular, one may expect companies to implement an RTO mandate because of pre-existing issues with employee retention, leading to endogeneity concerns. In order to estimate the counterfactual distribution, we use the distributional synthetic controls estimator from [Gunsilius \(2023\)](#). The idea is to extend the classical synthetic controls method ([Abadie and Gardeazabal, 2003](#); [Abadie et al., 2010](#); [Abadie, 2021](#)) to settings where the policy change of interest is implemented at an aggregate level, but the researcher has access to more granular data. This is exactly our setting, where the level of the intervention is at the company level, but the heterogeneity of the treatment is of fundamental interest. The distributional synthetic controls estimator is ideally suited for this setting as it estimates the heterogeneous treatment effect for the entire company and not an individual treatment effect for each unique worker. In particular, the method does not make the common assumption that individuals can be tracked over time. Relaxing this assumption allows for changes in the composition of workers due to both entry and exit into the company, which is a crucial aspect of within-firm employee distributions.

We briefly review the idea behind distributional synthetic controls. For more detailed information we refer to [Gunsilius \(2023\)](#) and the R implementation ([Van Dijcke et al., 2024](#)). Denote by $F_{jt,N}(y)$ the cumulative distribution function of the outcome of interest Y —in our case seniority or tenure—for company j at time t in the absence of the intervention. The distribution in the presence of intervention is denoted by $F_{jt,I}$. We write t^* for the time point of the implementation of the policy. By construction, the observed distribution F_{jt} coincides with $F_{jt,N}$ for all $t \leq t^*$. We denote the corresponding quantile function by

$$F_{jt}^{-1}(q) = \inf_{y \in \mathbb{R}} \{F_{jt}(y) \geq q\}, \quad q \in [0, 1].$$

In our setting the distributions F_{jt} are estimated from individual level data $Y_{i(t)jt}$, where $i(t)$ is the individual i at company j at time t . Individual observations are only used to estimate the distributions \hat{F}_{jtn} and corresponding quantiles \hat{F}_{jtn}^{-1} , where n is the number of cross-sectional observations in time period t , which is allowed to vary with both t and j .

In the synthetic controls setting we pick a target company $j = 0$ that at a point T^* implements the RTO mandate. We focus on Microsoft as our target company but replicate our main results for SpaceX and Apple. We only use companies as potential controls that had not implemented a form of RTO six months after the end of the sample period, or equivalently nine months after the target company’s RTO. Our goal is to estimate the counterfactual distribution $F_{0t,N}$ for time periods $t > T^*$ to see what would have happened to the outcome distribution of the target company had it not implemented the RTO mandate.

We have two main outcomes of interest, tenure at the company in number of days and seniority in terms of an employee’s title. We treat tenure as a continuous variable. Therefore, we estimate the counterfactual tenure distribution by obtaining the optimal weights $\vec{\lambda}_t^* = (\lambda_{1t}, \dots, \lambda_{Jt})$ in each time period $t \leq t^*$ by solving

$$\vec{\lambda}_t^* = \operatorname{argmin}_{\lambda \in \Delta^J} \int_0^1 \left| \sum_{j=1}^J \lambda_{jt} F_{jt}^{-1}(q) - F_{0t}^{-1}(q) \right|^2 dq,$$

forming one set of time-averaged weights $\vec{\lambda}^* = \frac{1}{t^*} \sum_{t=0}^{t^*} \vec{\lambda}_t^*$, and constructing the counterfac-

tual quantile function for the time periods $t > t^*$ via

$$F_{0t,N}^{-1}(q) = \sum_{j=1}^J \lambda_j F_{jt}^{-1}(q).$$

Here, $\Delta^J = \left\{ (\lambda_{1t}, \dots, \lambda_{Jt}) \in \mathbb{R}^J : \lambda_{1t}, \dots, \lambda_{Jt} \geq 0, \sum_{j=1}^J \lambda_j = 1 \right\}$ is the probability unit simplex in \mathbb{R}^J . The other outcome variable is seniority, which is a discrete variable with 10 ordinal values. Since the outcome is ordinal and not cardinal, it makes sense to use mixtures of distributions instead of mixtures of quantiles, as distributions preserve the ordinal structure; therefore, in line with the arguments in [Gunsilius \(2023\)](#), we compute the optimal weights $\vec{\lambda}_t^*$ in this case via

$$(2) \quad \vec{\lambda}_t^* = \operatorname{argmin}_{\lambda \in \Delta^J} \int_{\mathbb{R}} \left| \sum_{j=1}^J \lambda_{jt} F_{jt}(y) - F_{0t}(y) \right| dy,$$

and construct the counterfactual distribution in $t > t^*$ via

$$F_{0t,N}(y) = \sum_{j=1}^J \lambda_j F_{jt}(y).$$

Additionally, we also consider the share of leavers with a certain characteristic.

To obtain confidence intervals, we rely on the bootstrap. The bootstrap algorithm we use as well as conditions for its uniform convergence are provided in [Appendix B](#). We report uniform confidence bands throughout to account for simultaneous inference.

Finally, we perform inference via a permutation test, analogous to the classical setting ([Abadie et al., 2010](#)). To account for suboptimal fits in the replications of the distributions in the pre-treatment periods, we augment the proposed test in [Gunsilius \(2023\)](#) and construct the permutation test as the ratio of post-to-pre-intervention distances following [Abadie et al. \(2010\)](#); [Abadie \(2021\)](#) by computing the statistic

$$(3) \quad r_j = \frac{R_j(T^* + 1, T)}{R_j(1, T^*)}$$

when treating company j as the target and calculating the p -value for the permutation test as

$$p = \frac{1}{J+1} \sum_{j=0}^J H(r_j - r_0),$$

where $H(x)$ is the Heaviside function, which takes the value 1 if $x \geq 0$ and 0 if $x < 0$. Here,

$$R_j(t_1, t_2) = \left(\frac{1}{t_2 - t_1 + 1} \sum_{t=t_1}^{t_2} d_t^2 \right)^{1/2}$$

is the root mean squared prediction error in the distance d_t , which we take to be the 2-Wasserstein distance

$$d_t^2 = \int_{q_{\min}}^{q_{\max}} \left| \sum_{j=1}^J \lambda_j^* F_{jt}^{-1}(q) - F_{0t}^{-1}(q) \right|^2 dq$$

Table 3: Top 10 Synthetic Controls Weights: Microsoft

Company	Weight	Company	Weight
Amazon	0.1925	Protiviti	0.3185
Autodesk	0.1282	Linkedin	0.2054
Dell Technologies	0.1098	Intuit	0.1753
Slalom Consulting	0.0884	Dell Technologies	0.1547
Motorola Solutions	0.0859	Cisco	0.0567
3m	0.0805	Deloitte	0.0469
Cisco	0.0674	Intel Corporation	0.0422
DocuSign	0.0497	Hp	0.0003
Citrix	0.0364	Broadridge	0.0000
Cox Automotive Inc.	0.0328	Nvidia	0.0000

(a) Tenure

(b) Titles

Note: table shows names of 10 companies with largest synthetic controls weights for Microsoft’s RTO results shown in Figures 4 and 5.

between the predicted distribution using the optimal weights λ_j^* and the target. q_{\min} and q_{\max} are set equal to 0 and 1, respectively, for permutation tests on the full distribution, but can be restricted to a smaller interval to test inference on parts of the distribution as well.

3 Results

3.1 The causal effects of an RTO on tenure and seniority

Using the method described above, we estimate synthetic counterfactual tenure and title distributions for Microsoft. For the tenure distribution, we restrict the distribution to the bottom 90 quantiles as the top 10 quantiles are very sparse, leading to unstable estimates. In Table 3, we report the top 10 largest weights each of the control units’ distributions receive, together with the name of the control unit. Most of the companies receiving the largest weights are highly similar to Microsoft, being large diversified tech firms that produce both hardware and software, such as Amazon, Dell, Motorola, and Cisco.

With these weights in hand, we construct a distributional synthetic control for Microsoft. We find that the implementation of a return-to-office mandate in April 2022 led to significant employee outflows at Microsoft, with more senior staff leaving at higher rates. Figure 4 depicts this through the counterfactual changes in employee tenure.

For the left panel, the y-axis captures the difference between the observed distribution of tenure in number of days and the counterfactual one, estimated using the distributional synthetic controls method. The x-axis represents quantiles of these distributions. The synthetic control can accurately reconstruct the pre-RTO tenure distribution, as the difference between the two distributions is close to and not significantly different from zero at all quantile grid points. Following the return-to-office (RTO) mandate, we estimate a reduction in tenure among Microsoft’s employees that escalates with tenure duration, and only becomes

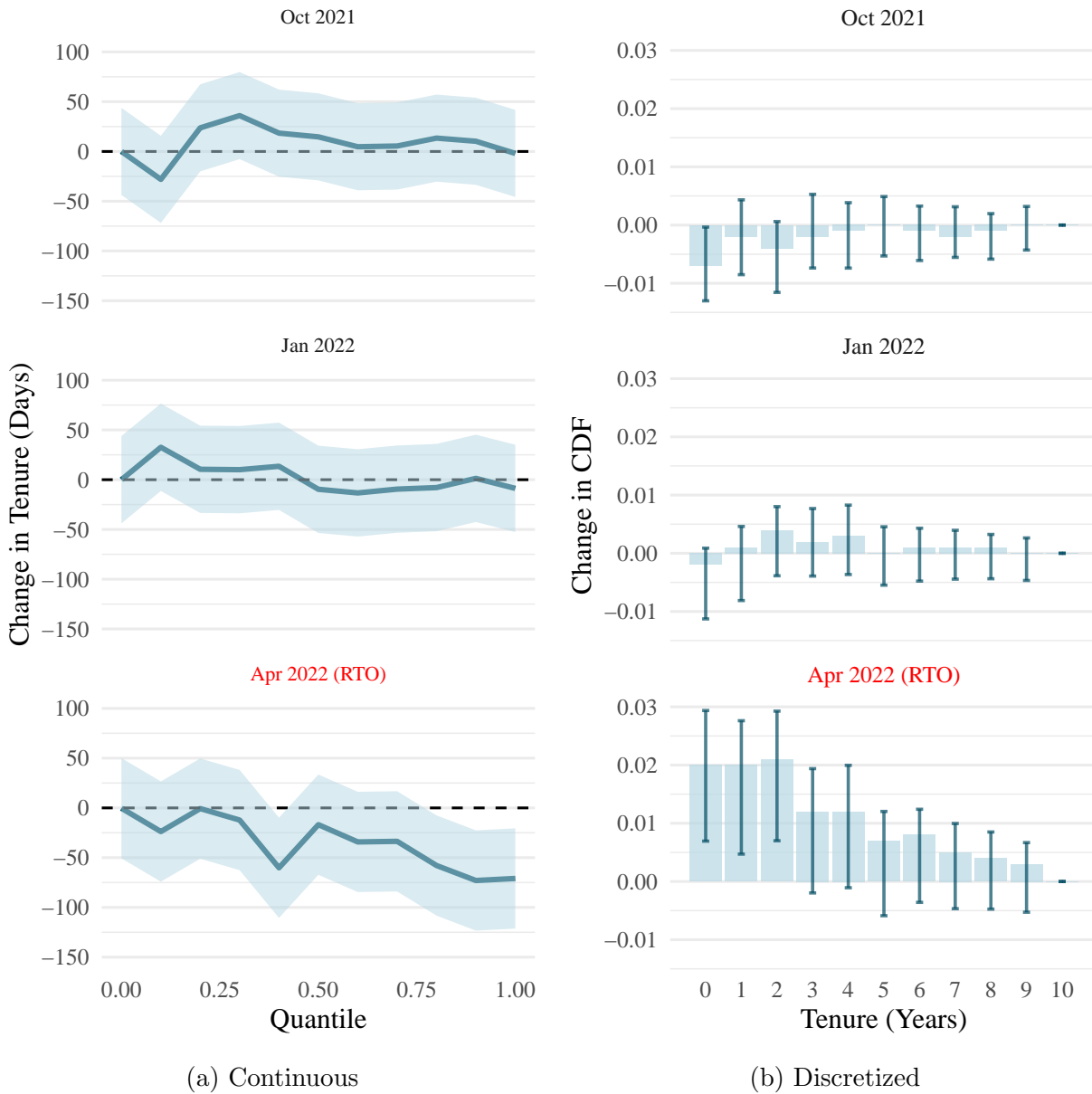


Figure 4: Tenure at Microsoft After RTO

Note: **(a)**: counterfactual change in quantile functions of tenure in days at Microsoft for two quarters before its RTO and one quarter after; **(b)** counterfactual change in CDFs of tenure in years for the same. Facet titles indicate first month of respective quarters. Estimated change is depicted by dark blue solid line and 95% bootstrapped uniform confidence intervals (see Appendix B) by light blue shaded area (1,000 bootstraps). Permutation test p-values are (see (3)): left panel 0.03, right panel 0.06. $N=1,141,518$, $J+1=32$.

statistically significant for the top 2 deciles. With pointwise confidence bands, this becomes the top 5 deciles (not reported). In terms of magnitudes, tenure at Microsoft decreased by two months at the top deciles. Our permutation test supports the validity of this inference, with a p-value of 0.03.

For an alternative perspective, we discretize the tenure measure into years and estimate the counterfactual change in the CDF of this discrete variable using the mixture of distributions approach in (2). Thus, we can interpret the y-axis as capturing the increase in probability mass at a given year of tenure. The resulting estimates are shown in the left panel of Figure 4. We replicate the pre-treatment distributions nearly perfectly. After treatment, we estimate a 6% increase in the probability mass of employees that have 0 to 3 years of tenure, corresponding to a relative decrease in the share of employees with longer tenure. Again, the permutation p-value of 0.06 supports the validity of the inference.

In Figure 5, we consider a more substantively different measure of the employee distribution: seniority. The figure depicts an increase in the mass of workers at lower-ranking titles at Microsoft after its RTO, using the mixture of distributions approach in (2). We estimate a statistically significant increase in the mass of employees in training, entry, and manager positions of over 4%. The estimator replicates the pre-treatment distribution well, with the confidence intervals including zero everywhere. We obtain nearly identical results when restricting the distribution to only the senior level and below, where most of the mass is concentrated. The permutation test once again supports the validity of the inference, with a p-value of 0.03.

We replicate these results almost identically for Apple and SpaceX in Figure 8. The result for the continuous tenure distribution does not replicate, likely because we observe fewer resumes at these companies, which makes it harder to accurately estimate the continuous tenure distribution. This explanation is supported by the bad pre-treatment fit of the tenure distribution for these companies. However, when we discretize it into 5-year bins, we estimate an increase in the mass of workers with less than 5 years of tenure of 1.5%. To replicate the result for the title distribution, we focus on the titles below the senior level at these companies, as there is very little mass at the higher titles.² Interestingly, the magnitude of the results appears to increase with the stringency of the RTO, with Apple seeing a leftward shift in mass of under 4% and no significant effects for managers, while up to 15% of the SpaceX distribution shifts leftward. This correlates with these companies mandating 1 and 5 days of office attendance each week, respectively, compared to Microsoft’s 50%.

Several points merit further discussion. First, the distribution of some of our control units may be affected by spillover effects, if one of the employees leaving Microsoft takes up employment at a control unit. We do not believe this will meaningfully bias our estimates, as only 5% of employees leave a company in a given quarter on average. Of those, we typically see only a handful of these leaving for the same, usually large, firm. Furthermore, we restricted our control sample to companies with at least 5% of the number of employees observed at Microsoft. Together, these facts imply that the spillover effects from control companies absorbing employees leaving the treated company are negligible relative to the effects on Microsoft.

Second, there could be anticipation effects that bias the magnitude of our estimates. For

²The results for Microsoft replicate identically when restricting the distribution in this way.

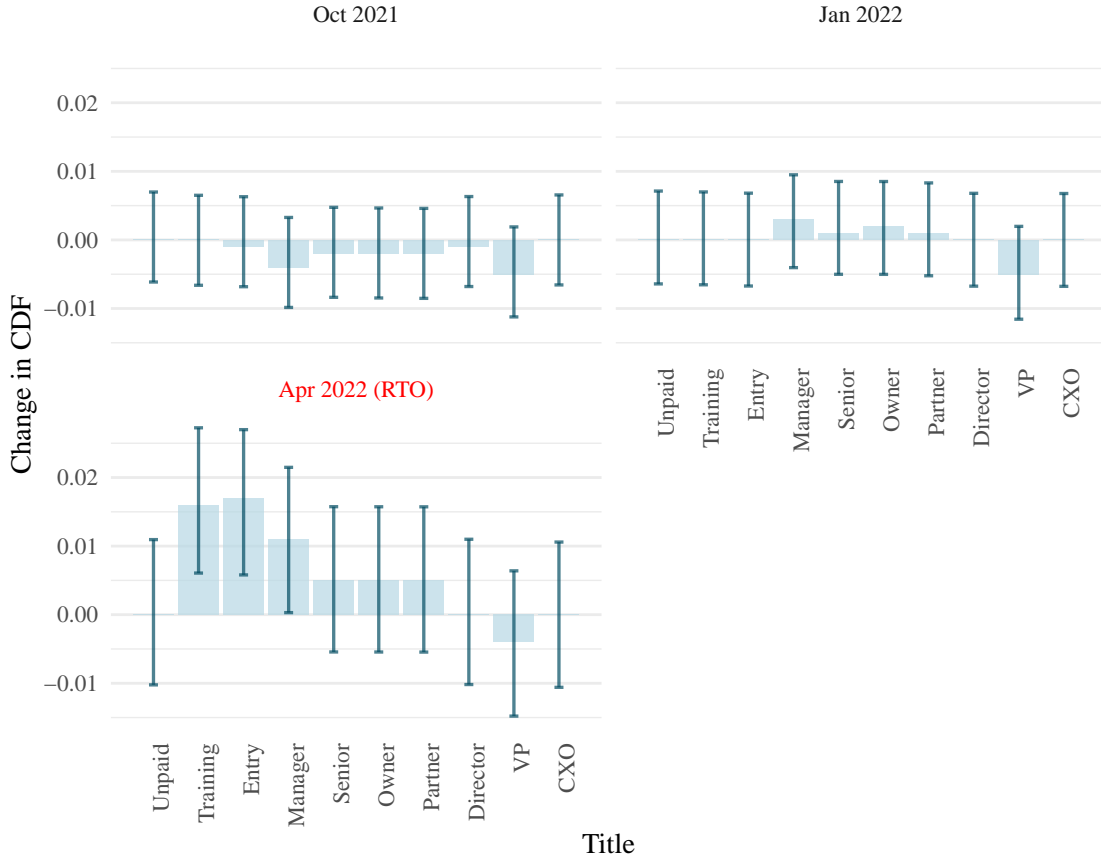


Figure 5: Titles at Microsoft After RTO

Note: figure shows counterfactual change in cumulative density functions of title distribution at Microsoft for two quarters before its return to office (RTO) and one quarter after. Facet titles indicate first month of respective quarters. Estimated change is depicted by light blue solid bar and 95% bootstrapped uniform confidence intervals (see Appendix B) by dark blue whiskers (1,000 bootstraps). Permutation test p-value (see (3)) is 0.03. $N=1,141,518$; $J+1=32$.

all three companies we consider, however, the RTO mandates were announced less than 2 months prior to their implementation.³ Given that a typical job search takes around 4.5 months on average (Bureau of Labor Statistics, 2024) and resignations usually only take effect at the end of the month, this left essentially no room for anticipation that is not captured by our research design. Of course, employees may have heard rumors before the official announcement. Both Microsoft and Apple, however, delayed several prior RTOs indefinitely due to pushback from employees and pandemic conditions (O’Loughlin, 2023, 2022). Survey evidence from May 2022 indeed suggested that such pushback to office attendance require-

³Microsoft announced plans for a phased return to the office in early February 2022 with mandatory office presence from April 2022 onwards (Tilley and Cutter, 2022). Apple announced its RTO mandate, which took effect in April 2022, in early March (Haring, 2022); while SpaceX’s RTO was announced with immediate effect via an internal email from Elon Musk to employees (Mac, 2022).

ments had been common and effective (Bloom, 2022). Thus, it seems likely that employees only acted on the mandates once they were actually implemented due to the uncertainty surrounding them as well as their previous successes in pushing back the dates. This does not contradict the possibility that employees planning to leave had already started their job search before the actual RTO, which would explain the sizeable tenure effects we observe only 3 months after implementation.

Third, due to the interference of layoff dates with RTO mandates, we are only able to study these mandates' effects one quarter after their implementation. It is possible that firms try to mitigate the shift in their tenure distribution after an RTO by increased hiring or internal promotions. This does not negate the significant costs associated with this shift and any mitigation thereof, which we discuss in more detail below.

3.2 *Where Do Employees Go After an RTO?*

Our main results provide robust evidence that long-tenured and senior employees depart from Microsoft at higher rates after it implements a return to office mandate. A natural follow-up question to ask is: where do these senior employees go? To answer this question, we construct a set of distributional synthetic controls for the employees who leave Microsoft in a given quarter. Since the leavers comprise a far smaller sample than the total employees, we focus on replicating an aggregate distributional statistic (Abadie et al., 2010), the share of leavers who have characteristic X, rather than the entire distribution.

This way, we find that Microsoft's RTO led to a counterfactual decrease in the share of its employees departing for startups, as shown in Figure 6. In particular, we proxy for whether a firm is a start-up in two ways: whether it has less than 50 employees, and whether it is below the 5th quantile in terms of its industry labor share. For both measures, we estimate an increase in the mass of workers leaving to non-startups of around 5% that is significant at the 95% level. Moreover, the permutation test for both measures indicates the inference is (marginally) valid, with p-values of 0.03 and 0.09, respectively. Combined with our main finding that more senior employees are leaving at higher rates, this result confirms the descriptive evidence in Figure 3 that the RTO caused a departure of senior employees to direct, large competitors. Thus, Microsoft's RTO appears to not only have had important human capital costs due to the hiring and training costs required to replace the senior employees lost, but also potential competitive side effects due to these employees taking operational knowledge with them.

Finally, in Figure 9 we further test for, but fail to find any evidence of, distributional differences in: the share of men vs. women leaving Microsoft; the share of leavers that flow into unemployment; the share of leavers that accept a demotion in their title at their new job; and the share of leavers changing roles at their new job. The lack of gender differences is interesting insofar as it suggests women are not acting on their stronger preference for working from home that was documented in several surveys (YouGov, 2022; Pelta, 2021). The lack of counterfactual changes in unemployment status, demotions, and job roles suggests that the senior employees leaving Microsoft have good outside options, in the sense that they appear to find employment in similar roles and levels of seniority. This aligns with the fact that Microsoft implemented its RTO mandate much earlier than most tech companies, and we can expect the value of employees' outside option to shrink as more companies implement

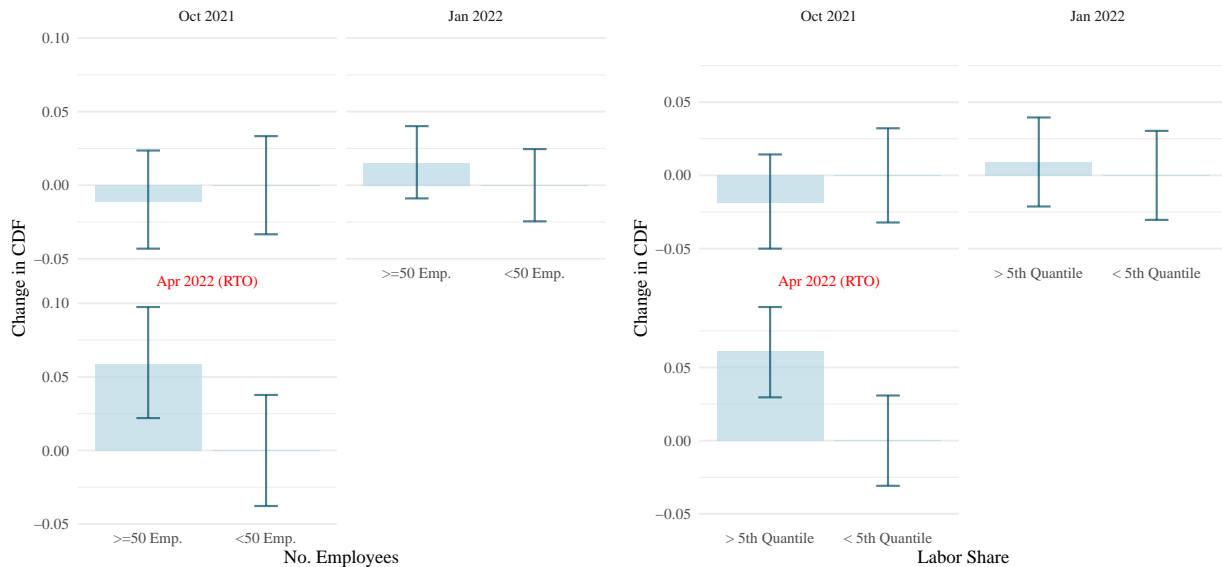


Figure 6: Share of Employees Leaving to Startups

Note: figure shows counterfactual change in share of employees leaving Microsoft by size of the company they leave to, measured by number of employees (larger or smaller than 50, left panel) and the company’s labor share (number of employees as a share of total industry employment, right panel). Estimated change is depicted by light blue solid bar and 95% bootstrapped uniform confidence intervals (see Appendix B) by dark blue whiskers (1,000 bootstraps). Pre-treatment fits are depicted in the top two bar plots, while the post-treatment counterfactual is shown in the bottom box plot. Permutation p-values (see (3)) are 0.03 and 0.09, respectively. $N=1,846$, $J+1=32$.

RTOs. At the same time, this finding also suggests that the share of companies in an industry that have returned to office will determine the negative impact of an RTO mandate on the company that implements it. That is because, all else equal, we can expect that the outflow of senior employees that we have documented will be larger when these employees’ outside options are better. This hints at the existence of a tipping point where companies implement RTO mandates at increasing rates when the value of employees’ outside options has diminished sufficiently. As a counterpoint, many companies started explicitly coding the degree of hybrid work allowed into new joiners’ contracts, which may generate path dependence both for fully remote companies as well as those that returned to the office. Exploring these dynamics could be a fruitful avenue for future research.

3.3 Discussion

This article estimates that a return to office mandate at Microsoft led to a significant outflow of senior employees to large competitors. What does this suggest about the larger ramifications of such mandates for the companies that implement them? To start, consider the reasons for implementing an RTO mandate. Managers tend to believe work-from-home decreases productivity (Bloom, Barrero, Davis, Meyer and Mihaylov, 2023b) and diffusion of knowledge within the firm (Barrero et al., 2023a), and face difficulties monitoring employ-

ees in remote-work settings (Microsoft, 2022). The academic literature provides support for some, but not all, of these concerns – see Barrero et al. (2023a) for an overview. On the other hand, employee satisfaction and retention have been found to decrease with a return to office (Bloom et al., 2015). Our results suggest that such retention issues may be more serious than previously thought, as senior employees and, indeed, the (senior) *managers themselves* are leaving. Thus, a firm returning to the office with the aim of increasing productivity or innovation may end up doing so in a narrow sense, improving the output metrics of those who stay, but harming them for the company as a whole.

In support of this, the academic literature provides ample evidence on the value of senior and long-tenured employees to the companies they work at. In broad terms, human capital has been theorized to be a key factor determining firms’ competitive advantage and estimated to be strongly related to firm performance (Crook et al., 2011). More specifically, worker output has been found to be monotonically increasing in tenure (Shaw and Lazear, 2008), with bosses estimated to be almost two times as productive as the average worker (Lazear et al., 2015). These effects are driven by (costly) employee training (De Grip and Sauermann, 2012; Konings and Vanormelingen, 2015) and the more general accrual of firm-specific human capital with tenure (Becker, 1962). Moreover, senior employees tend to not just *be* productive; they also strongly affect a firm’s wider productivity. Managers and management practices have been found to explain a large share of the variation in firm productivity (Bender et al., 2018; Metcalfe et al., 2023; Fenizia, 2022), especially in combination with higher-quality human capital (Bender et al., 2018). Employees working in research and development (R&D) not only increase a firm’s innovation rates, but also take their accumulated knowledge with them when they move to competitors (Palomeras and Melero, 2010; Cassiman and Veugelers, 2006). Relatedly, lower employee turnover has been found to increase firm investment and decrease new firm entry (Jeffers, 2024). Replacing skilled employees also incurs substantial hiring costs, equal to around 2–4 months of pay, that increase with the skill level (Blatter et al., 2012).

Taken together, this large body of literature implies that increased attrition of senior employees can substantially impede firm output, productivity, innovation, and competitiveness. This provides an important perspective on the implications of the return to office for the broader functioning of a company. Moreover, the downstream outcomes of the “uneven attrition” we document may be a fruitful avenue for future research.

4 Conclusion

The revolution in the workplace that was brought on by the COVID-19 pandemic continues to reverberate across offices globally, even as the pandemic itself retreats into history. While some companies have returned to the office, many others have gone fully remote. This paper has provided novel evidence on the implications of this bifurcation for the distribution of employees across and within companies. In particular, we provide causal evidence that three of the largest US tech companies – Microsoft, SpaceX, and Apple – faced a significant outflow of employees after implementing an RTO mandate, with more senior employees leaving at higher rates. We do this using a distributional extension of the classical synthetic controls estimator (Abadie and Gardeazabal, 2003; Gunsilius, 2023), for which we provide

conditions for the uniform validity of the bootstrap. Moreover, we provide both descriptive and causal evidence that these senior employees left for larger companies at higher rates than usual. Taken together, our findings imply that return to office mandates can imply significant human capital costs in terms of output, productivity, innovation, and competitiveness for the companies that implement them. Furthermore, the sorting of skilled labor across companies that our findings imply can have important consequences for the employment landscape, which we leave to future work.

References

- Abadie, A. (2021), ‘Using synthetic controls: Feasibility, data requirements, and methodological aspects’, *Journal of Economic Literature* **59**(2), 391–425.
- Abadie, A., Diamond, A. and Hainmueller, J. (2010), ‘Synthetic control methods for comparative case studies: Estimating the effect of california’s tobacco control program’, *Journal of the American statistical Association* **105**(490), 493–505.
- Abadie, A. and Gardeazabal, J. (2003), ‘The economic costs of conflict: A case study of the basque country’, *American economic review* **93**(1), 113–132.
- Aliprantis, C. D. and Border, K. C. (2006), *Infinite Dimensional Analysis: A Hitchhiker’s Guide*, Springer Science & Business Media.
- Barrero, J. M., Bloom, N. and Davis, S. J. (2021), ‘Why Working from Home Will Stick’.
URL: <https://www.nber.org/papers/w28731>
- Barrero, J. M., Bloom, N. and Davis, S. J. (2023a), ‘The Evolution of Work from Home’, *Journal of Economic Perspectives* **37**(4), 23–49.
URL: <https://pubs.aeaweb.org/doi/10.1257/jep.37.4.23>
- Barrero, J. M., Bloom, N. and Davis, S. J. (2023b), ‘The evolution of work from home’, *Journal of Economic Perspectives* **37**(4), 23–49.
- Barrero, J. M., Bloom, N., Davis, S. J., Meyer, B. H. and Mihaylov, E. (2022), ‘The Shift to Remote Work Lessens Wage-Growth Pressures’.
URL: <https://www.nber.org/papers/w30197>
- Becker, G. S. (1962), ‘Investment in human capital: A theoretical analysis’, *Journal of political economy* **70**(5, Part 2), 9–49.
- Bender, S., Bloom, N., Card, D., Van Reenen, J. and Wolter, S. (2018), ‘Management practices, workforce selection, and productivity’, *Journal of Labor Economics* **36**(S1), S371–S409.
- Blatter, M., Muehlemann, S. and Schenker, S. (2012), ‘The costs of hiring skilled workers’, *European Economic Review* **56**(1), 20–35.
- Bloom, N. (2022), ‘The great resistance: Getting employees back to the office’, *Stanford Institute for Economic Policy Research*.
- Bloom, N., Barrero, J. M., Davis, S. J., Meyer, B. and Mihaylov, E. (2023a), ‘Survey: Remote work isn’t going away—and executives know it’, *Harvard Business Review* **28**.
- Bloom, N., Barrero, J. M., Davis, S., Meyer, B. and Mihaylov, E. (2023b), ‘Research: Where managers and employees disagree about remote work’.
URL: <https://hbr.org/2023/01/research-where-managers-and-employees-disagree-about-remote-work>

- Bloom, N., Han, R. and Liang, J. (2023), How hybrid working from home works out, Technical report, National Bureau of economic research.
- Bloom, N., Liang, J., Roberts, J. and Ying, Z. J. (2015), ‘Does Working from Home Work? Evidence from a Chinese Experiment *’, *The Quarterly Journal of Economics* **130**(1), 165–218.
URL: <https://doi.org/10.1093/qje/qju032>
- Bureau of Labor Statistics (2024), ‘Average duration of unemployment in the united states from 1990 to 2023 (in weeks) [graph]’, <https://www.statista.com/statistics/217837/average-duration-of-unemployment-in-the-in-the-us/>.
- Bureau of Labor Statistics and NCWIT (2021), ‘Percentage of employed women in computing-related occupations in the united states from 2007 to 2020, by ethnicity’, [Graph]. In Statista. Retrieved April 19, 2024.
URL: <https://www.statista.com/statistics/311967/us-women-computer-workers-ethnicity/>
- Cassiman, B. and Veugelers, R. (2006), ‘In search of complementarity in innovation strategy: Internal r&d and external knowledge acquisition’, *Management science* **52**(1), 68–82.
- Clarke, F. H. (1990), *Optimization and nonsmooth analysis*, SIAM.
- Crook, T. R., Todd, S. Y., Combs, J. G., Woehr, D. J. and Ketchen Jr, D. J. (2011), ‘Does human capital matter? a meta-analysis of the relationship between human capital and firm performance.’, *Journal of applied psychology* **96**(3), 443.
- Dafermos, S. (1988), ‘Sensitivity analysis in variational inequalities’, *Mathematics of Operations Research* **13**(3), 421–434.
- De Grip, A. and Sauermann, J. (2012), ‘The effects of training on own and co-worker productivity: Evidence from a field experiment’, *The Economic Journal* **122**(560), 376–399.
- Ding, Y. and Ma, M. S. (2023), ‘Return-to-Office Mandates’.
URL: <https://papers.ssrn.com/abstract=4675401>
- Eeckhout, J. and Kircher, P. (2018), ‘Assortative matching with large firms’, *Econometrica* **86**(1), 85–132.
- Emanuel, N., Harrington, E. and Pallais, A. (2023), The power of proximity to coworkers: training for tomorrow or productivity today?, Technical report, National Bureau of Economic Research.
- Fenzia, A. (2022), ‘Managers and productivity in the public sector’, *Econometrica* **90**(3), 1063–1084.
- Gerhart, B. and Feng, J. (2021), ‘The resource-based view of the firm, human resources, and human capital: Progress and prospects’, *Journal of management* **47**(7), 1796–1819.
- Gunsilius, F. F. (2023), ‘Distributional synthetic controls’, *Econometrica* **91**(3), 1105–1117.
- Hansen, S., Lambert, P. J., Bloom, N., Davis, S. J., Sadun, R. and Taska, B. (2023), ‘Remote Work across Jobs, Companies, and Space’.
URL: <https://www.nber.org/papers/w31007>
- Haring, B. (2022), ‘Apple, google, twitter set back-to-work policies for their remote employees’, *Deadline* .
URL: <https://deadline.com/2022/03/apple-google-twitter-set-return-to-work-policies-1234972162/>
- Jeffers, J. S. (2024), ‘The impact of restricting labor mobility on corporate investment and entrepreneurship’, *The Review of Financial Studies* **37**(1), 1–44.

- Konings, J. and Vanormelingen, S. (2015), ‘The impact of training on productivity and wages: firm-level evidence’, *Review of Economics and Statistics* **97**(2), 485–497.
- Lazear, E. P., Shaw, K. L. and Stanton, C. T. (2015), ‘The value of bosses’, *Journal of Labor Economics* **33**(4), 823–861.
- Lise, J. and Robin, J.-M. (2017), ‘The macrodynamics of sorting between workers and firms’, *American Economic Review* **107**(4), 1104–1135.
- Mac, R. (2022), ‘Elon musk to workers: Spend 40 hours in the office, or else’, *The New York Times* .
URL: <https://www.nytimes.com/2022/06/01/technology/elon-musk-tesla-space-x-office.html>
- Mas, A. and Pallais, A. (2017), ‘Valuing alternative work arrangements’, *American Economic Review* **107**(12), 3722–3759.
- Metcalfe, R. D., Sollaci, A. B. and Syverson, C. (2023), Managers and productivity in retail, Technical report, National Bureau of Economic Research.
- Microsoft (2022), ‘Hybrid work is just work. are we doing it wrong?’. Accessed: 2024-04-26.
URL: <https://www.microsoft.com/en-us/worklab/work-trend-index/hybrid-work-is-just-work/>
- Mischke, J., Luby, R., Vickery, B., Woetzel, J., White, O., Sanghvi, A., Rhee, J., Fu, A., Palter, R., Dua, A. and Smit, S. (2023), Empty spaces and hybrid places: The pandemic’s lasting impact on real estate, Report, McKinsey & Company.
URL: <https://www.mckinsey.com/industries/real-estate/our-insights>
- O’Loughlin, H. (2022), ‘Microsoft’s return to office policy & timeline’, <https://buildremote.co/return-to-office/microsoft/>. Accessed: date-of-access.
- O’Loughlin, H. (2023), ‘Apple’s return to office policy & timeline’, <https://buildremote.co/return-to-office/apple/>.
- Palomeras, N. and Melero, E. (2010), ‘Markets for inventors: Learning-by-hiring as a driver of mobility’, *Management Science* **56**(5), 881–895.
- Parikh, N. and Boyd, S. (2014), ‘Proximal algorithms’, *Foundations and trends® in Optimization* **1**(3), 127–239.
- Pelta, R. (2021), ‘Survey: Men & women experience remote work differently’.
URL: <https://www.flexjobs.com/blog/post/men-women-experience-remote-work-survey/>
- Preiss, D. (2014), ‘Gâteaux differentiability of cone-monotone and pointwise lipschitz functions’, *Israel Journal of Mathematics* **203**(1), 501–534.
- Scoop (2023), The flex report: Tech industry deep dive, Technical report, Scoop.
URL: <https://www.flex.scoopforwork.com/reports/tech-industry>
- Shaw, K. and Lazear, E. P. (2008), ‘Tenure and output’, *Labour Economics* **15**(4), 704–723.
- Tilley, A. and Cutter, C. (2022), ‘Microsoft to start bringing employees back to the office’.
URL: <https://www.wsj.com/articles/microsoft-to-start-bringing-employees-back-to-the-office-11644865875>
- Unispace (2023), Returning for good, Global workplace report, Unispace.
URL: <https://www.unispace.com/the-reluctant-returner/download-english>
- van der Vaart, A. W. and Wellner, J. A. (1996), *Weak convergence and empirical processes*, Springer.
- Van Dijcke, D., Gunsilius, F. and He, S. (2024), *DiSCos: Distributional Synthetic Controls Estimation*. <http://www.davidvandijcke.com/DiSCos>.

van Triest, S. (2023), 'Your Boss Wants You at the Office: Supervisor Discretion and Employee Working from Home'.

URL: <https://papers.ssrn.com/abstract=4479293>

YouGov (2022), 'Working from home is working for working women', YouGov.

URL: <https://today.yougov.com/economy/articles/44482-working-home-working-for-working-women-poll>

A Additional Results

A.1 Figures

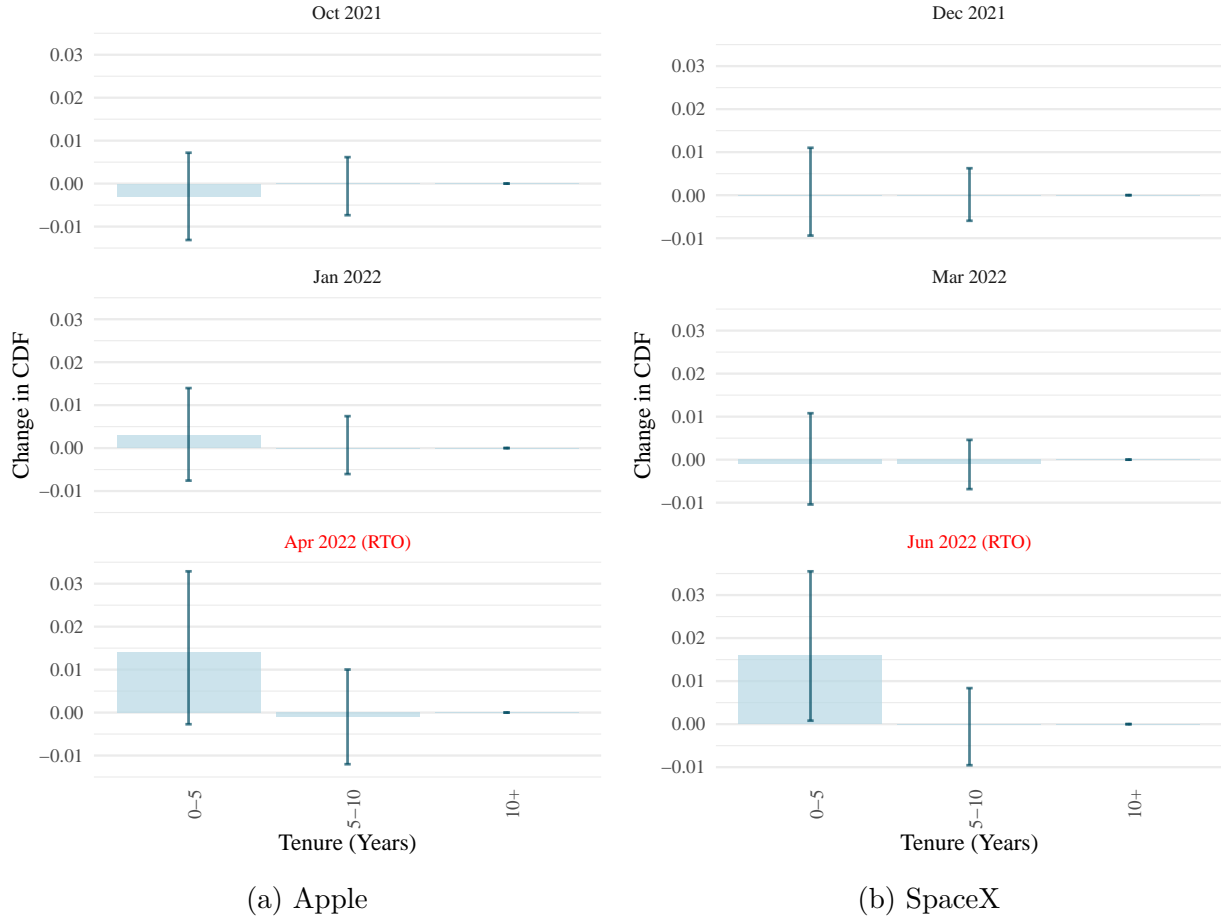


Figure 7: Tenure After RTO, Robustness

Note: figure shows counterfactual change in cumulative density functions of seniority levels at SpaceX and Microsoft for two quarters before its return to office (RTO) and one quarter after. Estimated change is depicted by light blue solid bar and 95% bootstrapped uniform confidence intervals (see Appendix B) by dark blue whiskers (1,000 bootstraps). Permutation test p-values (see (3)) are: Apple 0.25, SpaceX 0.04. Apple: $N=1,080,159$, $J+1=44$; SpaceX: $N=935,060$, $J+1=220$.

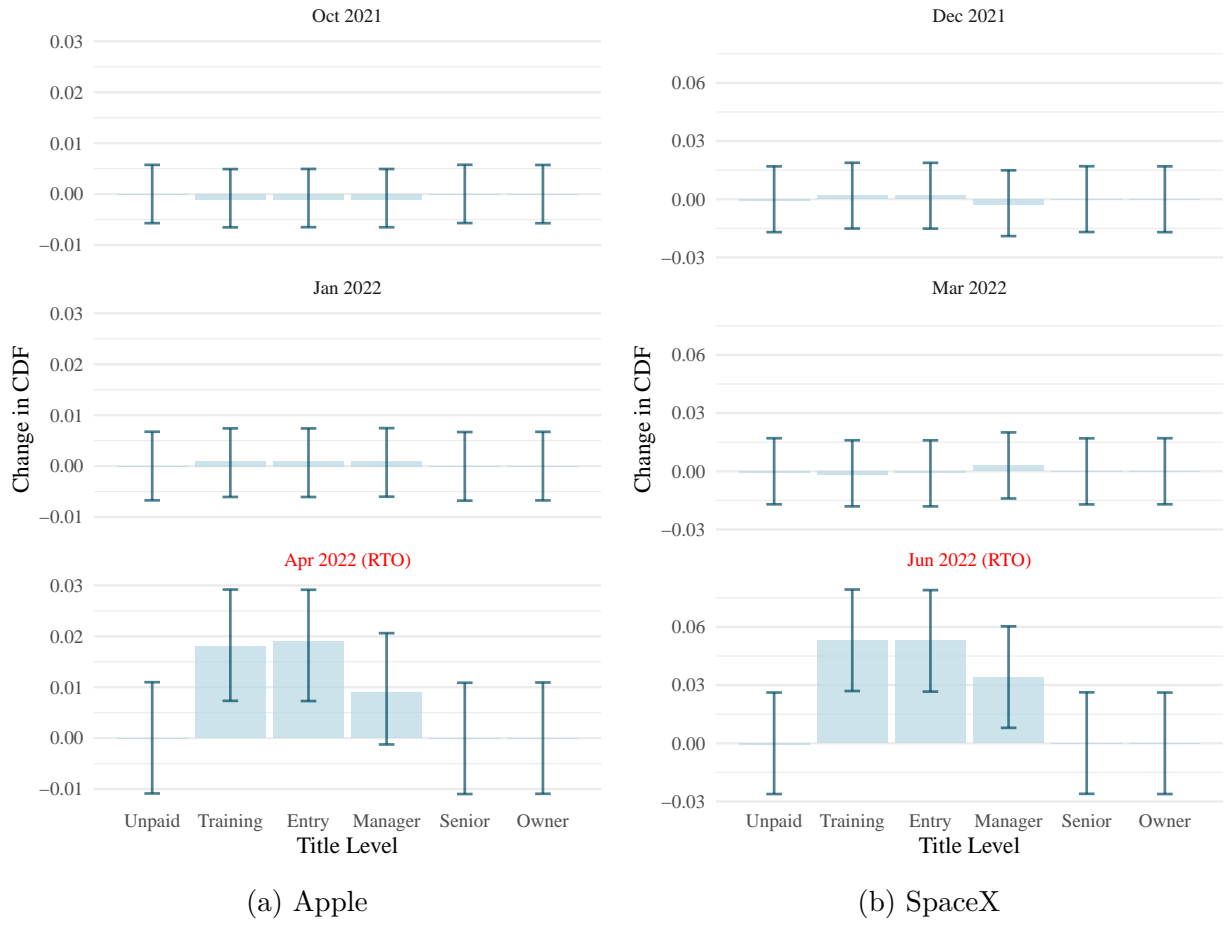


Figure 8: Titles After RTO, Robustness

Note: Estimated change is depicted by light blue solid bar and 95% bootstrapped uniform confidence intervals (see Appendix B) by dark blue whiskers (1,000 bootstraps). Permutation test p-values (see (3)) are: Apple 0.07, SpaceX 0.05. Apple: N=1,080,159, J+1=44; SpaceX: N=935,060, J+1=220.

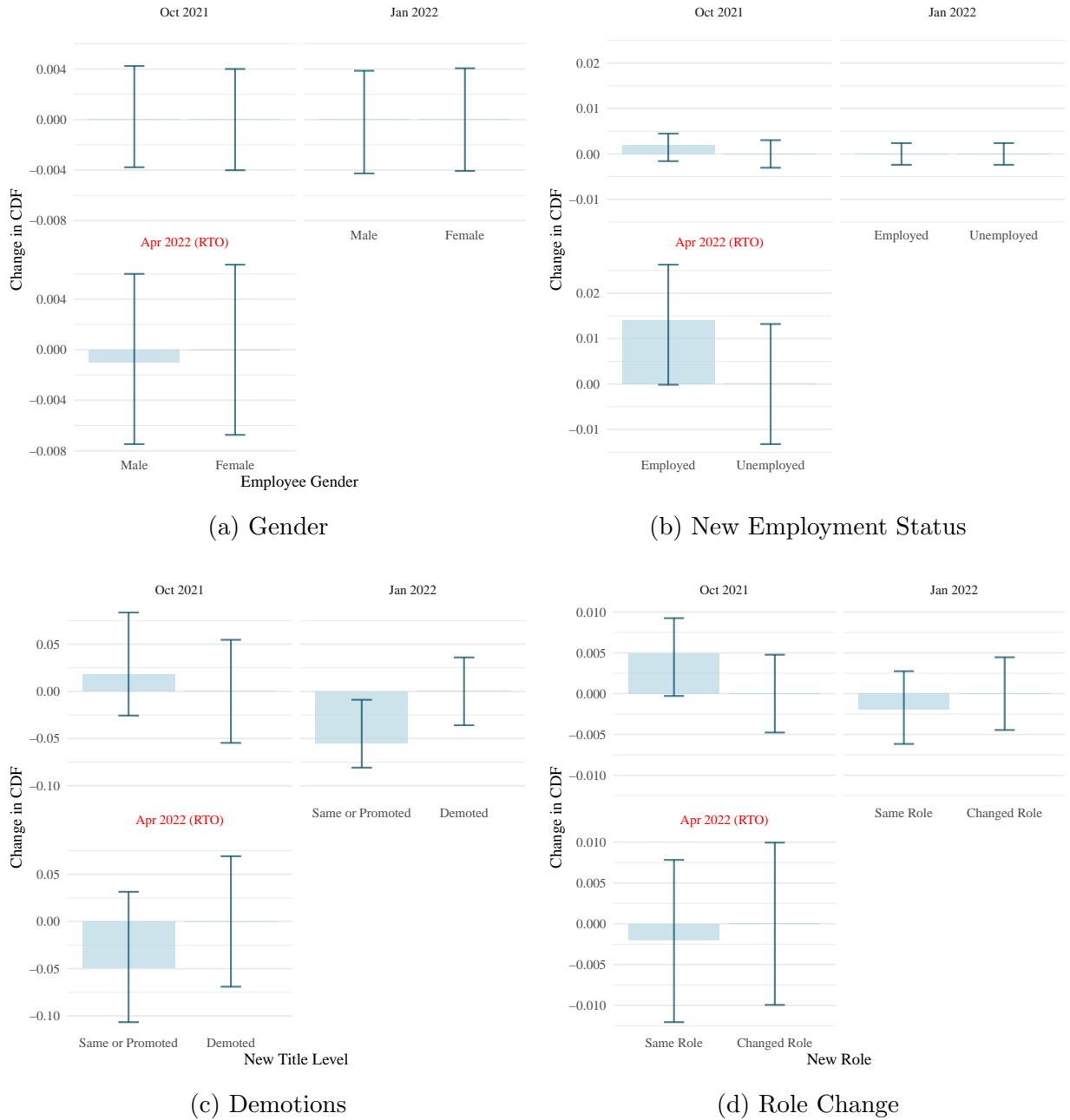


Figure 9: Gender, Employment, Demotions, and Role Changes Among Employees Leaving Microsoft

Note: Figure shows counterfactual change in share of employees leaving Microsoft by gender, new employment status, whether they accepted a demotion in their new job, and whether they changed roles in their new job. Estimated change is depicted by light blue solid bar and 95% bootstrapped uniform confidence intervals (see Appendix B) by dark blue whiskers (1,000 bootstraps). Pre-treatment fits are depicted in the top two bar plots, while the post-treatment counterfactual is shown in the third bar plot in each graph. Permutation test p-values (see (3)) are: gender 0.44, employment status 0.41, demotions 0.69, role change 0.88.

B Bootstrapping confidence intervals for the constructed quantile functions

In this section we provide assumptions on the data under which a canonical bootstrap procedure is valid for obtaining statistical uniform confidence intervals on the constructed quantile functions

$$(4) \quad F_{0t,N}^{-1} := \sum_{j=1}^J \lambda_j^* F_{jt}^{-1},$$

where λ_j^* are the optimal weights (Gunsilius, 2023). In the case where $t \leq t^*$, this is the “replicated” quantile function of the observable F_{0t}^{-1} ; in the case $t > t^*$, it is the counterfactual quantile function. The bootstrap procedure is captured in Algorithm 1.

In the following, to simplify notation, we assume that for all elements $j = 0, \dots, J$ and for all time periods t we have the same amount of observations $n_{jt} = n$. This restriction serves only to save on notation and the results can be extended to the case where we allow for different number of observations n_{jt} as long as they are asymptotically balanced, i.e., as long as $\frac{n_{ks}}{n_{jt}} \rightarrow \eta_{jkst}$ with $0 < \eta_{jkst} < +\infty$ for all $j, k \in \{0, \dots, J\}$ and all $s, t \leq T$.

Formally, validity of the bootstrap means that the bootstrapped empirical process

$$\tilde{\mathbb{G}}_n = \sqrt{n} \left(\tilde{F}_{0tn,N}^{-1} - \hat{F}_{0tn,N}^{-1} \right) = \sqrt{n} \left(\sum_{j=1}^J \tilde{\lambda}_{jn}^* \tilde{F}_{jtn}^{-1} - \sum_{j=1}^J \hat{\lambda}_{jn}^* \hat{F}_{jtn}^{-1} \right),$$

converges weakly to the standard empirical process

$$\mathbb{G}_n = \sqrt{n} \left(\hat{F}_{0tn,N}^{-1} - F_{0tn}^{-1} \right) = \sqrt{n} \left(\sum_{j=1}^J \hat{\lambda}_{jn}^* \hat{F}_{jtn}^{-1} - \sum_{j=1}^J \lambda_j^* F_{jtn}^{-1} \right),$$

where $\tilde{\lambda}_{jn}^*$ and \tilde{F}_{jtn}^{-1} are the bootstrapped optimal weights and quantile functions; $\hat{\lambda}_{jn}^*$ and \hat{F}_{jtn}^{-1} are the estimated optimal weights and quantile functions; and λ_j^* and F_{jt}^{-1} are the optimal weights and quantile functions in the population.

It is important to note that we can only show that the bootstrap is valid up to a null set in the Banach space $\times_{j=0}^J L^2([a, b])$ equipped with the Euclidean product topology. Since we work in infinite dimensional spaces, this null set is different from the classical “almost everywhere” statements, as an analogue to Lebesgue measure does not exist in infinite dimensional spaces. We rely on existing results showing that this is a null set in the sense that points of non-differentiability are in the σ -ideal $\tilde{\mathcal{C}}$ generated by sets “regularly null on curves running in the direction of some open cone” (Preiss, 2014, Definition 1.1). For practical purposes this “almost everywhere” specification is not a restriction, as the data-generating processes that can make the bootstrap invalid will not occur. We require the following assumption on the data generating process, the first part of which is a standard assumption for bootstrap validity of quantile estimators (van der Vaart and Wellner, 1996, Lemma 3.9.23).

Assumption 1. *For all $j = 0, \dots, J$ and all time periods $t \leq T$, we have the same number of observations n and all observations Y_{ijt} are iid draws from the distribution F_{jt} . In addition, the data-generating process is such that the corresponding distribution functions F_{jt} either*

Algorithm 1 The Bootstrap for Distributional Synthetic Controls

- 1: **Input:** Outcome data $\left\{ (Y_{ijt})_{i=1, \dots, N_{jt}} \right\}_{j=0, \dots, J; t=1, \dots, T}$, quantile grid $\mathcal{Q} := \left\{ \frac{q}{G} : q = 0, \dots, G \right\}$, empirical quantile functions $\hat{\mathbf{F}}_{tn}^{-1} := \left(\hat{F}_{jtn}^{-1}(q) : q \in \mathcal{Q} \right)_{j=0, \dots, J}$, index set $\mathcal{I}_{jt} := \{0, \dots, N_{jt}\}$, number of bootstrap draws B , significance level $\alpha \in (0, 1)$, time-specific weights estimators $\hat{\lambda}_{tn}^* \left(\hat{\mathbf{F}}_{tn}^{-1} \right) := \left(\hat{\lambda}_{jtn}^* \left(\hat{F}_{jtn}^{-1} \right) \right)_{j=1, \dots, J}$, the derived average weights $\hat{\lambda}_n^* := (\hat{\lambda}_{1n}^*, \dots, \hat{\lambda}_{Jn}^*) := \frac{1}{t^*} \sum_{t=1}^{t^*} \hat{\lambda}_t^*$, and counterfactual quantile estimator $\hat{F}_{0tn, N}^{-1} := \sum_{j=1}^J \hat{\lambda}_{jn}^* \hat{F}_{jtn}^{-1}$
- 2: **Output:** Confidence bands for the average weights $\hat{\lambda}_n^*$, counterfactual quantile function $\left\{ \hat{F}_{0tn, N}^{-1} \right\}_{t=1, \dots, T}$, and counterfactual quantile differences $\left\{ \hat{F}_{0tn}^{-1} - \hat{F}_{0tn, N}^{-1} \right\}_{t=1, \dots, T}$.
- 3: **for** $b = 1$ to B **do**
- 4: **for** $t = 1$ to T **do**
- 5: For each $j = 1, \dots, J$, draw a new sample $\tilde{\mathcal{I}}_{jt}^b$ of size N_{jt} without replacement from \mathcal{I}_{jt}
- 6: Estimate the new empirical quantile functions $\tilde{\mathbf{F}}_{tn}^{-1, b}$ from the resampled data $\left\{ (Y_{ijt})_{i \in \tilde{\mathcal{I}}_{jt}} \right\}_{j=0, \dots, J}$
- 7: Compute the optimal time-specific weights $\tilde{\lambda}_{tn}^{*, b} \left(\tilde{\mathbf{F}}_{tn}^{-1, b} \right)$ from the resampled quantile functions
- 8: **end for**
- 9: Calculate the resampled average weights $\tilde{\lambda}_n^{*, b}$ from the resampled time-specific weights
- 10: Construct the counterfactual quantile function $\left\{ \tilde{F}_{0tn, N}^{-1, b} \right\}_{t=1, \dots, T}$ and quantile differences $\left\{ \tilde{F}_{0tn}^{-1, b} - \tilde{F}_{0t}^{-1, b} \right\}_{t=1, \dots, T}$ from the resampled average weights and quantile functions
- 11: **end for**
- 12: Construct confidence intervals for each object of interest \mathbf{X} and its corresponding estimate $\hat{\mathbf{X}}$,

$$C(\alpha) := \left[\hat{\mathbf{X}} - t_{1-\alpha}, \hat{\mathbf{X}} + t_{1-\alpha} \right],$$

where $t_{1-\alpha}$ is the $1 - \alpha$ quantile of the maximum absolute difference between the bootstrapped estimates and the main estimate \mathbf{X}^b .

- (i) have compact support $[a, b]$ and are continuously differentiable with strictly positive density or
- (ii) have infinite support and for every $0 < q_1 < q_2 < 1$ there exists an $\varepsilon > 0$ such that F_{jt} is continuously differentiable on the interval $[a, b] \equiv [F_{jt}^{-1}(q_1) - \varepsilon, F_{jt}^{-1}(q_2) + \varepsilon]$ with positive density f .

Moreover, the $J \times J$ Hessian matrix \mathbf{H}_t with jk -th entry

$$\mathbf{H}_t(j, k) = \int F_{kt}^{-1}(q) F_{jt}^{-1}(q) dq$$

is strongly positive definite for all t in the sense that there exists $\alpha > 0$ such that

$$\langle \mathbf{H}_t x, x \rangle \geq \alpha \|x\|_2^2$$

for any $x \in \mathbb{R}^J$, where $\|\cdot\|_2$ denotes the Euclidean norm.

This assumption allows us to prove validity of the bootstrap for obtaining confidence intervals on the constructed quantile function; in the compact support case we can do this for the entire quantile range $[0, 1]$, while in the case of unbounded support we can only do it for any compact subinterval of the open interval $(0, 1)$. One can interpret the assumption on the matrix \mathbf{H}_t as essentially a ‘‘variance’’ condition. It requires that there is enough variation in the quantile functions and mirrors the classical variance condition for the invertibility of the design matrix in classical linear regression.

The following proposition is key to proving bootstrap validity. It proves Hadamard differentiability of the constructed $F_{0t,N}^{-1}$ based on the observable quantile functions F_{jt}^{-1} . Hadamard differentiability lets us apply the Delta method (van der Vaart and Wellner, 1996, chapter 3.9) to derive bootstrap validity. Since $F_{0t,N}^{-1}$ is a function of \mathbf{F}_t^{-1} , we work in the product spaces $\times_{j=0}^J L^2([a, b])$ and $\mathbf{C}[a, b] := \times_{j=0}^J C[a, b]$, which we equip with the standard Euclidean product norm. Note that we assume without loss of generality that all quantiles have the same support $[a, b]$; this can of course be relaxed at the cost of a significant increase in notational burden.

Proposition 1. *Under Assumption 1, the constructed quantile functions (4) are Hadamard differentiable at the inputs $\mathbf{F}_t := (F_{0t}, \dots, F_{Jt})$ tangentially to $\mathbf{C}[a, b]$, the J -dimensional product space of continuous functions on $[a, b]$, at $\tilde{\mathbf{C}}$ almost every \mathbf{F}_t^{-1} satisfying Assumption 1.*

Proof. Under Assumption 1 the quantile functions F_{jt}^{-1} are all Hadamard differentiable tangentially to $C[a, b]$ in the space $L^2([a, b])$ of square integrable functions with respect to Lebesgue measure (van der Vaart and Wellner, 1996, Lemma 3.9.23). Since $C[a, b]$ and $L^2([a, b])$ are Banach spaces, we can use the chain rule for Hadamard derivatives in Banach spaces (van der Vaart and Wellner, 1996, Lemma 3.9.3) and can therefore focus on the Hadamard differentiability of the optimal $\vec{\lambda}_t^*(\mathbf{F}_t^{-1})$ in $\mathbf{F}_t^{-1} = (F_{0t}^{-1}, \dots, F_{Jt}^{-1})$ in each time period t , obtained via

$$\vec{\lambda}_t^*(\mathbf{F}_t^{-1}) = \underset{\lambda \in \Delta^J}{\operatorname{argmin}} f(\vec{\lambda}_t, \mathbf{F}_t^{-1}) \equiv \underset{\lambda \in \Delta^J}{\operatorname{argmin}} \int_0^1 \left| \sum_{j=1}^J \lambda_{jt} F_{jt}^{-1}(q) - F_{0t}^{-1}(q) \right|^2 dq.$$

Step 1: Rewriting the constrained problem:

To show Hadamard differentiability Lebesgue almost everywhere, we use the fact that the optimization problem is strictly convex since the objective function is a squared L^2 -norm. This implies that the optimal solution $\vec{\lambda}_t^*$ is unique. Then we write the constrained problem as

$$\vec{\lambda}_t^*(\mathbf{F}_t^{-1}) = \operatorname{argmin}_{\lambda \in \mathbb{R}^J} \int_0^1 \left| \sum_{j=1}^J \lambda_{jt} F_{jt}^{-1}(q) - F_{0t}^{-1}(q) \right|^2 dq + \chi_{\Delta^J}(\vec{\lambda}),$$

where $\chi_A(x)$ is the indicator function that is zero if $x \in A$ and $+\infty$ otherwise.

The weights $\vec{\lambda}_t^*$ are optimal if and only if they satisfy the first order condition (Parikh and Boyd, 2014, p. 149)

$$0 \in \partial_{\vec{\lambda}} f(\vec{\lambda}_t^*, \mathbf{F}_t^{-1}) + \partial_{\vec{\lambda}_t^*} \chi_{\Delta^J}(\vec{\lambda}_t^*),$$

where $\partial_{\vec{\lambda}}$ denotes the vector subgradient with respect to $\vec{\lambda} \in \mathbb{R}^J$. This is equivalent to the fixed point condition (Parikh and Boyd, 2014, p. 150)

$$(5) \quad \vec{\lambda}_t^*(\mathbf{F}_t^{-1}) = \pi_{\Delta^J} \left(\vec{\lambda}_t^*(\mathbf{F}_t^{-1}) - \partial_{\vec{\lambda}} f(\vec{\lambda}_t^*, \mathbf{F}_t^{-1}) \right),$$

where π_{Δ^J} is the standard metric projection onto Δ^J using the Euclidean norm, which follows from e.g. Clarke (1990, Proposition 2.4.12).

Since f is differentiable in $\vec{\lambda}$ its subgradient $\partial_{\vec{\lambda}}$ is single-valued and coincides with the standard gradient. It is a $1 \times J$ -vector, with j -th entry computed via

$$\begin{aligned} & \partial_{\lambda_j} \int \left| \sum_{j=1}^J \lambda_{jt} F_{jt}^{-1}(q) - F_{0t}^{-1}(q) \right|^2 dq \\ &= \int \partial_{\lambda_j} \left| \sum_{j=1}^J \lambda_{jt} F_{jt}^{-1}(q) - F_{0t}^{-1}(q) \right|^2 dq \\ &= 2 \int \left(\sum_{k=1}^J \lambda_{kt} F_{kt}^{-1}(q) - F_{0t}^{-1}(q) \right) F_{jt}^{-1}(q) dq, \end{aligned}$$

where the second line follows from the dominated convergence theorem in conjunction with the boundedness of the quantile functions under Assumption 1.

We now want to take the Hadamard derivative $D_{\mathbf{F}_t^{-1}}$ of this first-order condition at an optimal $\vec{\lambda}_t^*$ with respect to the $J + 1$ -dimensional vector $\mathbf{F}_t^{-1} = (F_{0t}^{-1}, \dots, F_{Jt}^{-1})$. The projection π_{Δ^J} is only Lipschitz continuous in its argument (Aliprantis and Border, 2006, Lemma 6.54d) and hence by Rademacher's theorem only differentiable almost everywhere.

We therefore analyze the right-hand side of (5). Formally, using the chain rule of Hadamard derivatives in Banach spaces (van der Vaart and Wellner, 1996, Theorem 3.9.3), we have

$$\begin{aligned} & D_{\mathbf{F}_t^{-1}} \left(\vec{\lambda}_t^*(\mathbf{F}_t^{-1}) - \partial_{\vec{\lambda}} f(\vec{\lambda}_t^*, \mathbf{F}_t^{-1}) \right) \\ &= D_{\mathbf{F}_t^{-1}} \vec{\lambda}_t^*(\mathbf{F}_t^{-1}) - D_{\mathbf{F}_t^{-1}} \partial_{\vec{\lambda}} f(\vec{\lambda}_t^*, \mathbf{F}_t^{-1}) - \partial_{\vec{\lambda}}^2 f(\vec{\lambda}_t^*, \mathbf{F}_t^{-1}) D_{\mathbf{F}_t^{-1}} \vec{\lambda}_t^*(\mathbf{F}_t^{-1}), \end{aligned}$$

where $D_{\mathbf{F}_t^{-1}}\partial_{\vec{\lambda}}f(\vec{\lambda}_t, \mathbf{F}_t^{-1})$ is arranged as a Jacobian of dimension $J \times (J + 1)$, $\partial_{\vec{\lambda}}^2f(\vec{\lambda}_t, \mathbf{F}_t^{-1})$ is a Hessian of dimension $J \times J$ and $D_{\mathbf{F}_t^{-1}}\vec{\lambda}(\mathbf{F}_t^{-1})$ is again a Jacobian of dimension $J \times (J + 1)$.

We now compute the jk -th element in the Jacobian $D_{\mathbf{F}_t^{-1}}\partial_{\vec{\lambda}}f(\vec{\lambda}_t, \mathbf{F}_t^{-1})$ and the jk -th element in the Hessian $\partial_{\vec{\lambda}}^2f \equiv \mathbf{H}_t$. The Hadamard derivative of $\partial_{\lambda_j}f(\vec{\lambda}_t, \mathbf{F}_t^{-1})$ tangentially to $C[a, b]$ with respect to F_{kt}^{-1} for $k = 1, \dots, J$ is $\lambda_{kt} \int F_{kt}^{-1}(q)F_{jt}^{-1}(q) dq$, which follows from an application of the dominated convergence theorem in conjunction with the square integrability of the quantile functions and Hölder's inequality. For $k = 0$, the Hadamard derivative is $\int F_{0t}^{-1}(q)F_{jt}^{-1}(q) dq$. The jk -th entry in the Hessian $\partial_{\vec{\lambda}}^2f$ takes the form $\int F_{kt}^{-1}(q)F_{jt}^{-1}(q) dq$, again using the dominated convergence theorem.

The object of interest is $D_{\mathbf{F}_t^{-1}}\vec{\lambda}(\mathbf{F}_t^{-1})$, which is the Hadamard derivative of $\vec{\lambda}_t$ with respect to the quantile functions at the optimal value $\vec{\lambda}_t^*$. We want to use the fixed-point condition (5) to show that it is Hadamard differentiable up to a negligible set in $\mathbf{C}[a, b]$.

Step 2: Analyzing $D_{\mathbf{F}_t^{-1}}\vec{\lambda}_t(\mathbf{F}_t^{-1})$ at $\vec{\lambda}_t^$.*

To fix ideas, consider the case where the term inside the projection in (5) lies inside Δ^J . In this case the projection is the identity mapping and a rearranging of (5) gives

$$D_{\mathbf{F}_t^{-1}}\partial_{\vec{\lambda}}f(\vec{\lambda}_t, \mathbf{F}_t^{-1}) + \partial_{\vec{\lambda}}^2f(\vec{\lambda}_t, \mathbf{F}_t^{-1})D_{\mathbf{F}_t^{-1}}\vec{\lambda}(\mathbf{F}_t^{-1}) = 0.$$

Hence, $D_{\mathbf{F}_t^{-1}}\vec{\lambda}_t(\mathbf{F}_t^{-1})$ is Hadamard differentiable if $\partial_{\vec{\lambda}}^2f(\vec{\lambda}_t, \mathbf{F}_t^{-1}) \equiv \mathbf{H}_t$ is invertible.

We now consider the general case. We first show that $\vec{\lambda}_t^*(\mathbf{F}_t^{-1})$ defined via the fixed-point relation (5) is Lipschitz continuous in \mathbf{F}_t^{-1} (Dafermos, 1988). Then we rely on infinite-dimensional analogues of Rademacher's and Stepanov's theorem (Preiss, 2014) that guarantee that the optimal weights are Hadamard differentiable in \mathbf{F}_t^{-1} up to a small set in the Banach space $\times_{j=0}^J L^2([a, b])$.

The Lipschitz continuity of $\vec{\lambda}_t^*$ in \mathbf{F}_t^{-1} follows from the exact same argument as Lemma 2.4 in Dafermos (1988), but in a Banach space setting; since the proof of Lemma 2.4 only works with norms, this argument directly extends to the Banach space setting. We only have to check the corresponding (Lipschitz-) continuity conditions. For this, define $g(\vec{\lambda}_t^*, \mathbf{F}_t^{-1}) := \partial_{\vec{\lambda}}f(\vec{\lambda}_t^*, \mathbf{F}_t^{-1})$ and note that the restriction on \mathbf{H}_t implies that $g(\vec{\lambda}_t^*, \mathbf{F}_t^{-1})$ is strongly monotone in $\vec{\lambda}_t^*$ in the sense of equation (1.3) in Dafermos (1988). Indeed,

$$\begin{aligned} \langle g(\vec{\lambda}_t^*, \mathbf{F}_t^{-1}) - g(\tilde{\lambda}_t^*, \mathbf{F}_t^{-1}), \vec{\lambda}_t^* - \tilde{\lambda}_t^* \rangle &= 2 \sum_{j=1}^J \sum_{k=1}^J (\lambda_k - \tilde{\lambda}_k)(\lambda_j - \tilde{\lambda}_j) \int F_{kt}^{-1}(q)F_{jt}^{-1}(q) dq \\ &= 2(\vec{\lambda}_t^* - \tilde{\lambda}_t^*)\mathbf{H}_t(\vec{\lambda}_t^* - \tilde{\lambda}_t^*)^\top \\ &\geq \alpha \|\vec{\lambda}_t^* - \tilde{\lambda}_t^*\|_2^2 \end{aligned}$$

for some $\alpha > 0$.

Analogously, the Lipschitz continuity of g in $\vec{\lambda}_t^*$ for fixed \mathbf{F}_t^{-1} as required in equation (1.4) in Dafermos (1988) follows from

$$\left\| g(\vec{\lambda}_t^*, \mathbf{F}_t^{-1}) - g(\tilde{\lambda}_t^*, \mathbf{F}_t^{-1}) \right\|_2$$

$$\begin{aligned}
&= 2 \left[\sum_{j=1}^J \left(\sum_{k=1}^J (\lambda_{kt} - \tilde{\lambda}_{kt}) \int F_{kt}^{-1}(q) F_{jt}^{-1}(q) dq \right)^2 \right]^{1/2} \\
&\leq 2 \left[\sum_{j=1}^J \left(\sum_{k=1}^J (\lambda_{kt} - \tilde{\lambda}_{kt})^2 \right) \left(\sum_{k=1}^J \left(\int F_{kt}^{-1}(q) F_{jt}^{-1}(q) dq \right)^2 \right) \right]^{1/2} \\
&= 2 \left(\sum_{k=1}^J (\lambda_{kt} - \tilde{\lambda}_{kt})^2 \right)^{1/2} \left[\sum_{j=1}^J \sum_{k=1}^J \left(\int F_{kt}^{-1}(q) F_{jt}^{-1}(q) dq \right)^2 \right]^{1/2} \\
&= 2 \|\vec{\lambda}_t^* - \tilde{\lambda}_t^*\|_2 \|\mathbf{H}_t\|_F,
\end{aligned}$$

where the inequality follows from Cauchy-Schwarz and $\|\cdot\|_F$ denotes the Frobenius norm, so that the Lipschitz constant is $2\|\mathbf{H}_t\|_F$.

We also need to show that $g(\vec{\lambda}_t^*, \mathbf{F}_t^{-1})$ is Lipschitz continuous in \mathbf{F}_t^{-1} . Since \mathbf{F}_t^{-1} is an element in a product space, we can show Lipschitz continuity with respect to every entry separately. Indeed, denoting \mathbf{F}_{-kt} the vector where we fix every F_{jt}^{-1} except F_{kt}^{-1} , we write

$$\begin{aligned}
&\left\| g(\vec{\lambda}_t^*, \mathbf{F}_{-kt}) - g(\vec{\lambda}_t^*, \tilde{\mathbf{F}}_{-kt}) \right\|_2 \\
&= 2 \left[\sum_{j=1}^J \left(\lambda_{kt} \int (F_{kt}^{-1}(q) - \tilde{F}_{kt}^{-1}(q)) F_{jt}^{-1}(q) dq \right)^2 \right]^{1/2} \\
&\leq 2 \left[\sum_{j=1}^J \lambda_{kt}^2 \left(\int |(F_{kt}^{-1}(q) - \tilde{F}_{kt}^{-1}(q)) F_{jt}^{-1}(q)| dq \right)^2 \right]^{1/2} \\
&\leq 2 \left[\lambda_{kt}^2 \int (F_{kt}^{-1}(q) - \tilde{F}_{kt}^{-1}(q))^2 dq \sum_{j=1}^J \int (F_{jt}^{-1}(q))^2 dq \right]^{1/2} \\
&= 2 |\lambda_{kt}| \|F_{kt}^{-1} - \tilde{F}_{kt}^{-1}\|_{L^2([0,1])} \sqrt{\sum_{j=1}^J \|F_{jt}^{-1}\|_{L^2([0,1])}^2}
\end{aligned}$$

by Cauchy-Schwarz and Hölder's inequality, which by the fact that $F_{jt}^{-1} \in L^2([0,1])$ (which holds since all quantile functions are bounded by Assumption 1) for all j, t implies Lipschitz continuity in each F_{kt}^{-1} separately and hence in the product topology.

Finally, since Δ^J does not vary with \mathbf{F}_t^{-1} , the projection is trivially Lipschitz continuous with respect to \mathbf{F}_t^{-1} . Since all four criteria are satisfied in our case, the same argument as in Lemma 2.4 of [Dafermos \(1988\)](#) then implies that $\vec{\lambda}_t^*(\mathbf{F}_t^{-1})$ is Lipschitz continuous in \mathbf{F}_t^{-1} .

Based on Lipschitz continuity, we now use analogues of Rademacher's and Stepanov's theorem in Banach spaces to prove that $\vec{\lambda}_t^*(\mathbf{F}_t^{-1})$ is Hadamard differentiable tangentially to $\mathbf{C}[a, b]$ up to a negligible set. There are several results in the mathematical literature, but we rely on the results in [Preiss \(2014\)](#). In particular, since $\vec{\lambda}_t^*(\mathbf{F}_t^{-1})$ maps to \mathbb{R}^J , which trivially is a Banach space with the Radon-Nikodym property, Theorem 4.2 in [Preiss \(2014\)](#) shows that $\vec{\lambda}_t^*(\mathbf{F}_t^{-1})$ is Hadamard differentiable at $\tilde{\mathcal{C}}$ almost every \mathbf{F}_t^{-1} at which $\vec{\lambda}_t^*(\mathbf{F}_t^{-1})$ is Lipschitz

continuous.

Step 3: From Hadamard differentiability of $\vec{\lambda}_t^$ to $F_{0t,N}^{-1}$.* The above argument was for each t . Since the optimal $\vec{\lambda}^*$ used in the construction of $F_{0t,N}^{-1}$ is averaged over time periods $t \leq t^*$, this immediately implies that $\vec{\lambda}^*$ is Hadamard differentiable tangentially to $\mathbf{C}[a, b]$ at \mathbf{F}_t^{-1} at $\tilde{\mathcal{C}}$ almost every \mathbf{F}_t^{-1} . Applying the chain rule of Hadamard derivatives in form of the product rule once more (e.g. [van der Vaart and Wellner, 1996](#), section 3.9.4.4) then implies that $F_{0t,N}^{-1}$ is Hadamard differentiable at every t tangentially to $\mathbf{C}[a, b]$ at \mathbf{F}_t^{-1} at $\tilde{\mathcal{C}}$ almost every \mathbf{F}_t^{-1} satisfying Assumption 1. \square

Based on Proposition 1 we can then show that the bootstrap works almost everywhere, where the “almost everywhere” statement is with respect to $\tilde{\mathcal{C}}$ almost every \mathbf{F}_t^{-1} satisfying Assumption 1. In the following, we assume for the sake of simplicity that for all units j and time periods t we have the same number of observations $n_{jt} \equiv n$, which is only done to simply the statement. As mentioned earlier, if the number of observations are asymptotically equivalent, this result can be extended to this setting.

Theorem 1. *Suppose that $n_{jt} = n$ for all j, t and let Assumption 1 hold. Then the bootstrap works in probability at $\tilde{\mathcal{C}}$ -almost every \mathbf{F}_t^{-1} in the sense that*

$$\sup_{h \in \text{BL}} \left| \mathbb{E}_B h(\tilde{\mathbb{G}}_n) - \mathbb{E} h(\mathbb{G}_n) \right| \xrightarrow{P} 0.$$

In this result BL denotes the set of bounded Lipschitz functions, \mathbb{E}_B denotes the conditional expectation with respect to the data process (based on the number of resamples B_{inj} for every observation Y_{ijt}), “ \xrightarrow{P} ” denotes convergence in probability, and all probabilities and expectations are considered to be outer measures (e.g. [van der Vaart and Wellner, 1996](#), section 1.2). In words, this implies that asymptotically, the distributions of the empirical bootstrap process conditional on the data and the true empirical process are equivalent in probability, and that one can use the bootstrapped empirical process to obtain the confidence intervals.

Proof. The result follows immediately from the fact that under Assumption 1 the empirical quantile processes corresponding to F_{jt}^{-1} are bounded Donsker classes, which implies that the bootstrap procedure works for them in the sense written out in the theorem; this follows from Example 3.9.24 in [van der Vaart and Wellner \(1996\)](#). Then Proposition 1 implies the Hadamard differentiability of the constructed quantile functions at \mathbf{F}_t^{-1} at $\tilde{\mathcal{C}}$ almost every \mathbf{F}_t^{-1} . The conclusion then follows from the Delta method for bootstrap ([van der Vaart and Wellner, 1996](#), Theorem 3.9.11). \square