



ROCKWOOL Foundation Berlin

Institute for the Economy and the Future of Work (RFBerlin)

DISCUSSION PAPER SERIES

285/26

Equity and Efficiency in Anti-Discrimination Policy: Evidence from an Anonymous Hiring Pilot

Ohto Kanninen, Sanni Kiviholma, Tuomo Virkola

Equity and Efficiency in Anti-Discrimination Policy: Evidence from an Anonymous Hiring Pilot

Authors

Ohto Kanninen, Sanni Kiviholma, Tuomo Virkola

Reference

JEL Codes: J15, J24, J68, J71, M51

Keywords: anonymous hiring, discrimination, information frictions, job search

Recommended Citation: Ohto Kanninen, Sanni Kiviholma, Tuomo Virkola (2026): Equity and Efficiency in Anti-Discrimination Policy: Evidence from an Anonymous Hiring Pilot. RFBerlin Discussion Paper No. 285/26

Access

Papers can be downloaded free of charge from the RFBerlin website: <https://www.rfberlin.com/discussion-papers>

Discussion Papers of RFBerlin are indexed on RePEc: <https://ideas.repec.org/s/crm/wpaper.html>

Disclaimer

Opinions and views expressed in this paper are those of the author(s) and not those of RFBerlin. Research disseminated in this discussion paper series may include views on policy, but RFBerlin takes no institutional policy positions. RFBerlin is an independent research institute.

RFBerlin Discussion Papers often represent preliminary or incomplete work and have not been peer-reviewed. Citation and use of research disseminated in this series should take into account the provisional nature of the work. Discussion papers are shared to encourage feedback and foster academic discussion.

All materials were provided by the authors, who are responsible for proper attribution and rights clearance. While every effort has been made to ensure proper attribution and accuracy, should any issues arise regarding authorship, citation, or rights, please contact RFBerlin to request a correction.

These materials may not be used for the development or training of artificial intelligence systems.

Imprint

RFBerlin
ROCKWOOL Foundation Berlin –
Institute for the Economy
and the Future of Work

Gormannstrasse 22, 10119 Berlin
Tel: +49 (0) 151 143 444 67
E-mail: info@rfberlin.com
Web: www.rfberlin.com



Equity and Efficiency in Anti-Discrimination Policy: Evidence from an Anonymous Hiring Pilot*

Ohto Kanninen[†] Sanni Kiviholma[‡] Tuomo Virkola[§]

October 8, 2026

Abstract

Anonymous hiring restricts the information available to hiring managers, potentially creating a trade-off between equity and hiring efficiency. We leverage a pilot program by the City of Helsinki in which hiring for some, but not all, job titles was subject to anonymization. We find that applicants with immigrant background were more likely to be hired without detrimental effects on job turnover or wages. Announcing anonymization increased job applications, but not by groups that ex-post had the most to gain. Our results suggest anonymization can be an effective anti-discrimination policy at least when it's mandated and not voluntary for hiring managers.

JEL Codes: J15, J24, J68, J71, M51

Keywords: anonymous hiring, discrimination, information frictions, job search

*Acknowledgements: We thank participants at Society of Labor Economists conference, European Association for Labour Economists conference, Network of European Labour Market Research Institutes workshop, Australian Gender Economics Workshop, Finnish Economic Association Meeting, and Labore, VATT and Helsinki GSE seminars, as well as Michèle Belot, Cristina Bratu, Yajna Govind, Hannu Karhunen, Hanna Pesola, Matti Sarvimäki, and Marian Vidal-Fernandez for helpful comments. We also thank the City of Helsinki for providing us with ample information on the details of the pilot and access to the administrative data. We are grateful to the AMI Foundation for funding. Sanni Kiviholma received funding from the Finnish Cultural Foundation. Tuomo Virkola gratefully acknowledges financial support from the Research Council of Finland (# 360450). Any errors are ours.

[†]Chief Researcher at the Labour Institute for Economic Research & University Researcher at the University of Helsinki, Finland, email: ohto.kanninen@labore.fi.

[‡]Senior Researcher at the Labour Institute for Economic Research, Finland, email: sanni.kiviholma@labore.fi

[§]Academy Research Fellow at VATT Institute for Economic Research and Helsinki GSE, Finland, email: tuomo.virkola@vatt.fi.

1 Introduction

Job applicants whose names signal membership in disadvantaged groups often face a penalty in the hiring process (Riach and Rich 2002; Bertrand and Mullainathan 2004; Rich 2014; Zschirnt and Ruedin 2016; Kline et al. 2024). This has motivated the introduction of anonymous hiring policies that remove information about applicants’ names and group membership from the applications. Anonymization should make it harder for employers to directly discriminate in the early stages of the hiring process. However, limiting information about a job candidate might affect the quality of the hiring process if the name or group membership conveys some information about the candidate’s expected productivity or suitability for the job. Previous work shows that information is important for hiring decisions and wage trajectories (Farber and Gibbons 1996; Altonji and Pierret 2001; Pallais 2014; Carranza et al. 2022), but evidence on the costs and benefits of anonymization remain limited and could explain why it has not become common practice among employers.

To make progress in understanding these potential trade-offs, we study a public sector hiring pilot at the City of Helsinki, the largest employer in Finland. During the pilot, new hires for a pre-defined set of job titles were assigned to an anonymous hiring process in which hiring managers could not see identifying information about a job candidate before calls to interviews were sent out. This setting allows us to use a difference-in-difference framework comparing hiring across job titles over time with a plausibly exogenous change in anonymization. A key feature of the pilot was that it was mandated and hiring managers could not use their own discretion, which matters because voluntary anonymization could be undermined if the managers most prone to discriminate are the least willing to adopt it.

Another major advantage of the pilot is that, for the first time, we can leverage administrative data on the full hiring process as well as the post-hiring outcomes to empirically evaluate key policy trade-offs throughout the hiring pipeline. First, we use data on job

applicants, calls to interviews, and new hires. These data allow us to separately study each of the potentially important steps in the hiring process that may eventually lead to hiring a job candidate. Second, using payroll data on the employees, we are able to track employment and earnings of new hires and compare these outcomes for workers who were selected via an anonymous process with those who were not. This allows us to test how the hiring process affects the quality of matches between jobs and hires, which has not been possible before. Finally, we leverage a switch in the anonymization policy after the pilot and data on hiring manager characteristics to study how the take-up and effectiveness of anonymization depends on whether it is mandated or voluntary, an issue which can have important consequences for the policy's effectiveness.

We find that the City of Helsinki's anonymization policy increased the likelihood that applicants with immigrant background were hired. Our estimates suggest that applicants with a foreign name were 4.5 p.p. more likely to receive a call for an interview. While the name and identity of job candidates are revealed during the job interviews, the effect on hiring decisions remains positive at 3.1 p.p. It would be hard to explain this pattern with simple animus or taste-based discrimination, but it could be consistent with unconscious biases or biased beliefs about foreign-named applicants. Overall, the magnitudes suggest that names are important signals for managers in the hiring stage. In contrast, we do not observe that anonymization would have affected the gender composition of calls to interviews or hiring decisions.

While the policy was effective in promoting equity, we find no changes in the match quality. In particular, we see no differences in wages or worker turnover across hires made through the pilot and hires made in the business-as-usual format. We also see no change in the probability of filling vacancies which is an important margin for employers. This suggests that, while hiring managers do pay attention to the names of the job applicants, removing that information does not necessarily make hiring less efficient. However, we do observe that the policy slowed down the hiring process, by increasing

the time from vacancy posting to a worker starting on the job. We attribute this to the technical constraint that, to preserve anonymity, hiring managers could not interview applicants before the application deadline, preventing them from interviewing applicants as applications arrived.

We also find new evidence that anonymizing hiring processes and advertising this policy can affect the application decisions of potential job candidates. In particular, we find that anonymization increases the number of job applications. This increase is not driven by applications from individuals with foreign names who were targeted with the policy, and who ex-ante would have been expected to benefit the most. Instead, more Finnish-named applicants and women apply for these positions. These results are informative about how the content of job postings may attract different types of applicants, but they also have implications for the total effects of hiring individuals with different characteristics. In particular, our results suggest that the combination of changes in application decisions and the changes in hiring probabilities conditional on applying imply a positive total hiring effect for both foreign-named and women applicants, although with imprecisely estimated magnitudes.

Finally, we find that when anonymization is made voluntary for hiring managers, it leads to selective take-up and becomes less effective in leveling the playing field. First, during the pilot, when anonymization is mandated for pre-determined job titles, manager characteristics do not predict take-up as expected. After the pilot, however, manager characteristics become correlated with anonymization decisions. Second, comparing voluntary anonymizations to similar hires prior to the pilot, we find that they do not increase calls to interviews or hiring of applicants with immigrant background. These results suggest voluntary anonymizations are limited in their effectiveness. This plausibly explains why our baseline results differ from those estimated in the previous literature where anonymization is voluntary and the minority applicants face a hiring penalty (Behaghel et al. 2015).

Our study relates to a large literature on labor market discrimination which demonstrates that employers use names as informative signals about job applicants, at least in the early stages of recruitment (e.g., Lang and Lehmann (2012) and Neumark (2018)). Less is known about how the removal of identifying information affects hiring decisions and the quality of the hiring process. For example, if employers use names as signals of worker productivity, their removal might lead to employers using less informative alternative signals and this may not improve the outcomes of the targeted applicant group. Recent work on removing information about job applicants' criminal records ("ban-the-box") has provided some evidence in this direction (Agan and Starr 2018; Doleac and Hansen 2020). This literature also suggests that the effectiveness of such policies may depend on whether it is implemented in the public or private sector context (Shoag and Veuger 2021; Rose 2021; Jackson and Zhao 2017; Craigie 2020; Pyle 2024)

Our study also relates to research on anonymous hiring policies, which have shown mixed results (Goldin and Rouse 2000; Behaghel et al. 2015; Krause et al. 2012; Åslund and Skans 2012; Kolev et al. 2019; Neumark 2024). These studies are focused on the early recruitment stages, finding reduced discrimination in calls to interviews but not necessarily in final hiring outcomes. Our difference-in-difference research design improves on previous work that compares hiring outcomes of different applicant groups across two public sector administrative units in Sweden without exogenous or time-variation in the policy (Åslund and Skans 2012). On the other hand, Behaghel et al. (2015) suggest that minority applicants can be disadvantaged if anonymization is made voluntary and those most likely to discriminate opt out. Our design and results inform policy makers and employers about the effects of mandated anonymization where selective take-up has limited opportunities to undermine the policy. In addition, previous work has not addressed how anonymization affects job applications nor whether it affects the efficiency of the hiring process. We are for the first time able to analyze the entire recruitment process, from applications to hiring decisions, and analyzing the effects on job duration and wages on the

job. This allows us to better assess key trade-offs that are induced by the policy and which employers need to consider when making decisions about their recruitment processes.

Finally, our study relates to recent work on how different recruitment techniques and information in job advertisements affect job seekers' application decisions. For example, Belot et al. (2022), Banfi and Villena-Roldán (2019), and Kircher (2022) provide evidence that job seekers react to wage information. Recent work also suggests that explicit gender preferences (Kuhn and Shen 2013; Kuhn et al. 2020; Kuhn and Shen 2023; Card et al. 2023) and equality of opportunity or affirmative action statements (Leibbrandt and List 2018; Flory et al. 2021; Ibanez and Riener 2018) can affect job application decisions. When potential applicants observe that recruiting is anonymized, it may affect their beliefs about acceptance rates or quality of the workplace. We provide new evidence on how a specific anti-discrimination policy affects application decisions.

2 The City of Helsinki Anonymous Hiring Pilot

We study an anonymous hiring program at the City of Helsinki, a public sector employer, that aimed to reduce discrimination against job seekers with immigrant background. In December 2018, the City Council approved a motion to enable anonymous hiring processes to "advance the opportunities of individuals who face possible inadvertent discrimination."¹ This initiative was partly motivated by the under-representation of individuals with immigrant background among the City's employees, and by the existing evidence that immigrants are likely to face a penalty in the early stages of recruitment. Yet the implementation was more generally aiming to anonymize the hiring process along more than just the immigration dimension.

The pilot ran between February and December 2020. For the duration of the pilot, the City of Helsinki selected 13 job titles that would be subject to an anonymous hiring

¹The motion (in Finnish): <https://dev.hel.fi/maatokset/asia/hel-2018-006254/kvsto-2018-22/>

process. The number of applicants and vacancies by group is presented in Appendix Table F1. The anonymization was mandated and not dependent on the hiring manager's individual discretion. This creates the variation that we leverage in our empirical setting. However, take-up was imperfect as 44 percent of recruitment processes in the treated job titles were conducted anonymously during the pilot. This reflects i) initial randomization within the selected job titles which was stopped in May and ii) HR capacity constraints in setting up the randomization process rather than managerial discretion (Appendix A). The selected job titles included, e.g., administrative assistants and secretaries, youth workers, coordinators, and planners, and they were scattered across the city's administrative units. However, the pilot excluded job titles, such as nurses and practical nurses, for which the city faced hiring difficulties. As a result, the pilot is informative about hiring in contexts without significant labor scarcity, when recruiters are not compelled to hire unsuitable candidates.

During the pilot, all identifiable information about the candidate was hidden before invitations to job interviews were sent out. The HR department forwarded software-generated summary files of applicants to the hiring managers. These included information on the applicant's job experience, education, and degree titles, as well as a motivation letter.² In contrast, no information on the applicant's name, address, age, language skills, or educational institution was included to remove close proxies for group identity. The length of previous work experience was shown, but the exact dates were not. After sending out calls to interviews, full applicant files became available to hiring managers, and all interviews and hiring decisions were done in a standard, non-anonymous manner. Overall, this anonymization process may be expected to create more informational friction than simply removing applicants' names.

Anonymization was salient to job seekers and the city's employees. First, the policy

²The City asks applicants not to provide personal information in open text fields, and our manual audit of applications in Appendix D shows that the motivation letter and other open text fields did not systematically reveal applicants' background, and no more often in anonymized than in standard processes.

was publicly discussed in the media at the time of the City council’s decision in late 2018. This announcement could affect some potential job seekers’ decisions to apply for a job at the City of Helsinki. However, the implementation details, including the timing of pilot and treatment job titles, were not public knowledge at the time and were only decided shortly before the pilot started. Second, and more important for our analysis, all anonymized job postings in the pilot’s treatment group were advertised to job seekers as being subject to anonymization, with “Anonymous:” included in the vacancy title and an explanation provided in the job posting. In contrast, job postings in the control group and before the pilot include an equal-opportunity statement encouraging applicants from all backgrounds to apply. Finally, all managers observe when they are hiring anonymously and managers hiring to treatment and control groups may be aware of the policy’s broad targets. This means we cannot separately identify the effects of pure anonymization from the announcement of policy targets. Yet, our estimates are likely to reflect arguably more policy-relevant parameters when anonymization is accompanied with announced general policy goals.

Following the pilot, the City of Helsinki decided to make anonymization a voluntary recruitment tool for hiring managers. Thus, starting from January 2021, hiring managers could anonymize any hiring processes they were responsible for, regardless of job title. This policy change provides us an opportunity to understand the take-up of anonymization by hiring managers.

3 Data

We use administrative data from the City of Helsinki recruitment system that covers the universe of job openings and job applications between September 2019 and August 2023 (see Appendix A for details). We observe all basic information submitted by applicants to the recruitment system, including job experience, education, and language skills. Ad-

ditionally, it contains information from the advertisements as observed by the applicants, including the job title and whether the hiring process was anonymous. To measure the primary hiring outcomes, our data also contain information on which applicants were called to an interview and which applicants received a final job offer. Importantly, our data allow us to study hiring decisions with similar information that was available to hiring managers prior to sending out invitations for job interviews. Our sample concentrates on job ads with one vacancy.

We construct name-based measures of (perceived) immigration background and gender using publicly available information on Finnish first and last names from the Digital and Population Data Services Agency. First, we hired a research assistant to classify all names into either Nordic or non-Nordic sounding names. We then match this classification to job applicants and identify perceived immigrant background when the first name or the last name is classified as foreign. Second, to measure applicants' gender, we calculate the share of Finns who are women for each specific name. Most names are unique to either men or women, but when a name is used by both men and women, we classify it based on the majority gender.

To measure wages and job duration, we match all new hires to administrative payroll data using names and birth dates (see the Appendix B for details). The payroll data contain information on the monthly earnings at the individual level. We classify a worker as employed by the City of Helsinki whenever the monthly earnings are positive. We show in the Appendix A that there 1) remains significant variation in starting wages and promotions for similar jobs, and 2) temporary contracts and probation periods are important mechanisms through which job duration can adjust, suggesting that both can be informative about worker job-match quality.

4 Research Design

Our research design uses the variation in anonymized recruiting processes induced by the pilot program. Following the literature, we estimate how the probability of becoming interviewed or hired differs for applicants in job titles that were in the pilot and those that were not. Thus, in our baseline analysis, we leverage information at the applicant level, and the specification takes the form

$$Y_{ijt} = \alpha + \beta_g \text{Treated}_j \times \text{Post}_t + \delta X_{ijt} + \lambda_t + \theta_j + \varepsilon_{ijt}, \quad (1)$$

where Y_{ijt} is the outcome of applicant i who applied to a job of job title j at time t , Treated_j is a dummy for the job-title being assigned to anonymization during the pilot, Post_t is a dummy for the pilot duration and X_{ijt} is a set of (optional) controls including job and applicant level characteristics. We estimate the model separately for applicant groups g and test for differences in hiring outcomes across groups in an analogous triple-diff specification. To the extent that recruitment outcomes in the treated and control groups would have evolved in parallel in the absence of the pilot program and there are no anticipation effects, this specification yields an unbiased estimate of the intention to treat $\hat{\beta}_g$ of assignment on the treated openings. We view the no-anticipation assumption as plausible, as the specific job titles selected for the pilot were only decided and published shortly before the pilot began in 2020, and we may evaluate the plausibility of the parallel trends assumptions with pre-trends.

Our setting relies on variation in hiring processes between job titles so it is informative to look at how applications and hiring outcomes differ across job titles before the pilot started. Table 1 Column (1) shows the average characteristics of applicants and hires and Columns (2) and (3) show control and treated job titles separately. On average, there are some observable differences in age, education, and job experience among the applicants to these job openings. To test our identifying assumption, we divide the pre-treatment

period into two ten-week periods and study the evolution of covariates between these two periods. Column (4) shows the differences in pre-trends between treated job titles and control titles. Of the 21 pre-trends, three are significant at the 10 percent level (the share of temporary contracts, the share of foreign last names among applicants, and the share of recruits with a bachelor’s degree). This supports our main identifying assumption of parallel trends.

5 Results

This section first presents our main results on hiring decisions, and then discusses effects on job applications and match-quality, managers’ selection into anonymization and finally robustness of our results.

5.1 Hiring Decisions and Job Applications

In Figure 1 Panel A, we show event-study estimates on hiring separately for foreign- and Finnish-named applicants, and in Panel B we show the change in hiring gaps. Specifically, we replace the post indicator in our baseline model (Equation 1) with monthly event-time indicators. This allows us to transparently describe changes in our main outcome of interest around the policy introduction. The estimates show no statistically significant differential pre-trends in hiring decisions in the five months leading to the start of the pilot, providing support for our main identifying assumption. Second, after the start of the pilot, foreign-named applicants are more likely to be hired relative to Finnish-named applicants. The gap increases over time as the take-up of anonymization increases after the initial randomization within the treated job titles is dropped in May. In the Appendix Figure G2, we show that the estimates for calls to interviews are qualitatively similar and in favor of applicants with immigrant background.

Table 2 shows our main results on hiring decisions using the pooled diff-in-diff and

triple-difference models which allow us to gain more statistical power. First, Panel A Column (1) shows applicants with a foreign name are 4.5 p.p. more likely to be called to an interview when applying for a job that is part of the anonymous hiring pilot. The statistically insignificant point estimate for applicants with a Finnish name is -0.7 p.p. (Column 3). This results in a reduction of the interview gap by 5.2 p.p. relative to applicants with Finnish names (Column 5). The estimates change little if we add individual-level controls for demographics and skills (Columns 2, 4 and 6), which could be otherwise important determinants of hiring decisions. Our triple-diff design also allows to include vacancy fixed effects, which allows us to control for differences in the types of vacancies the two groups apply to, yielding very similar point estimates (Column 7).

Second, Columns (8)–(14), show estimates on hiring decisions that are positive but smaller in magnitude for foreign-named applicants (+3.1 p.p., Columns 8–9). However, the estimates for Finnish-named applicants become negative (-2.4 p.p., Columns 10–11), resulting in a similar change in the hiring gap between the groups. Again, we do not observe any meaningful changes in the estimates when we control for individual-level characteristics or vacancy fixed effects. This suggests that anonymization gives more interview opportunities for applicants with foreign names without a detectable cost for those with Finnish names, but as each opening has a single vacancy, hiring outcomes are largely zero-sum (Table 3, discussed below). Anonymization thus affects not only the early stage of the hiring process but also the final hiring decision. This would suggest that the literature on resume audits, which relies on the early stages of the hiring process, can be informative about final hiring decisions, a matter which has been brought up in previous literature (Cahuc et al. 2019).

Our estimates suggest that anonymization can have meaningful effects on hiring outcomes for immigrant job seekers. The estimated 4.5 p.p. interview effect is a 71 percent increase relative to the baseline (6.3 percent) and the 3.1 pp hiring effect roughly doubles the hiring rate from the baseline (3.1 percent). These translate into around 5 p.p. reduc-

tions in the gaps relative to Finnish-named applicants in interviews and hires. This is slightly larger than the baseline gap in hiring rates, which is consistent with immigrants applying to lower-level positions relative to their potential skills. These estimates are of the same order of magnitude but opposite in sign compared to Behaghel et al. (2015). They show that voluntary anonymization reduced minority interviews from 9.3 to 4.7 percent and increased majority interviews from 11.6 to 17.7 percent, resulting in a 10.7 p.p. widening of the interview gap, around twice our estimate. The estimated widening of the gap in hiring decisions is smaller at 3.7 p.p.

However, their estimates are based on a design with full take-up of anonymization while in our context take-up was partial and the estimates based on an ITT design. If we were to impose the exclusion restriction that the pilot did not affect hiring in processes that were not anonymized due to administrative frictions, the estimates on the gap scaled up by the take-up would be 12 and 11 percentage points, which still falls within the same absolute range as the negative estimates in Behaghel et al. (2015). However, as discussed in Section 2 and Appendix A, we cannot rule out that managers in the treatment group who did not hire anonymously were directly affected by the policy. Hence, we consider this as an upper bound. In Appendix Table F2, we also provide estimates that exclude the early implementation months during which only a subset of hirings in the treatment group were aimed to be anonymized, which raises both the take-up rate as well as the point estimates and the estimates fall within the ITT and this upper bound as expected.

The pilot also provides an opportunity to study the effects on gender discrimination. However, we do not observe significant changes in the gender composition of hires (Table 2, Panel B). The point estimates for interviews are negative for women relative to men but statistically insignificant, and the point estimates for hiring decisions are very close to zero. Event study estimates also confirm that there are no pre-trends in hiring by gender (Figure G2 Panels (e) and (f)). In the Appendix, we also show that the effects on hiring of immigrant men and immigrant women are very similar and positive (Ta-

ble F3). Anonymization may thus have limited impacts along dimensions where there may be limited discrimination at baseline (Chohlas-Wood et al. 2021). In Appendix E, we also study the impacts on hiring applicants age 50 or above and find positive effects on interviews but not on final hiring outcomes.

Application choices could be important for understanding the total effects of anonymous hiring policies on hiring outcomes. This is especially the case if the aim of the policy is to reduce potential discrimination but also to encourage potentially discriminated groups to apply. Table 3 Panels A and B show the effects on applications and the hiring process. First, job openings that were anonymized and advertised as anonymous received 5.7 more applicants (Panel A, Column 1), or around a 30 percent increase from a baseline of 19 applicants per job. Interestingly, this is almost all driven by Finnish-named applicants and not by applicants with immigrant background (Columns 2-3). Second, there is a similar but not statistically significant increase in the number of interviews (0.4, 29 percent from the baseline of 1.5, Panel A, Column 4). Third, there is no significant change in the number of job openings that lead to a hire, suggesting that anonymity does not necessarily make it harder to find suitable candidates (Column 5).

In Table 3, Panel B, we show how anonymization affects the types of job seekers who apply for jobs. As a result of the increase from Finnish-named applicants, the share of applicants with immigrant background decreases by 5.2 p.p. (16 percent). On the other hand, we observe that women's share increases by 5.5 p.p. (Column 2), showing that a large share of the increase comes from women. Anonymization could also plausibly affect applications and hiring decisions in other ways. For example, anonymization may shift the hiring manager's attention from applicants' group identity to their labor market skills and experience. Our estimates account for these compositional shifts by including worker-level controls and showing that the results remain similar independent of applicant skills. In Panel B columns (3)-(5), we show the impacts are positive for applicants education and negative for previous experience and age. In the Appendix Figure G6, we

show estimates on applications and applicant characteristics in a event-study design.

Now, the total effect of anonymization on hiring is a combination of changes in hiring probability among applicants and changes in the applicant composition. To measure the total effects that combine these two channels, Table 3 Panel C shows how the composition of hirees changes. We find a +4 p.p. effect on total share of hires with immigrant background. These estimates are now less precise relative to the estimated hiring gap in part because we may no longer take advantage of the individual-level variation and compare outcomes at the job opening level, but also because the two effects go in opposite directions. For the gender composition, the combination of an increase in applications and no changes in the hiring probabilities means that there is a larger share of women interviewed and hired under the anonymized processes. The point estimates for hiree skills, experience and age become negative but mostly not statistically significant at standard levels. We explore a related age-based channel in Appendix E, where we replicate our main analysis for applicants aged 50 and older. These results, however, illustrate the two potential mechanisms through which anonymization may affect final hiring outcomes, which employers may care about.

The results indicate that job applicants respond to the announcement of an anonymized hiring process. Women may perceive that their chances improve or that the anonymization sends a positive signal about job amenities. In contrast, we find no increase in applicants with immigrant background which could suggest potential misperception about expected hiring chances or some counteracting negative signals about workplace amenities.

5.2 Match Quality: Employment and Wages

To understand whether there is a potential efficiency cost to anonymous hiring, we study the quality of the job-worker match by comparing the post-hiring outcomes of workers hired in the treatment vs. control groups. Table 3 Panel D shows our difference-in-

difference estimates. First, Column 1 shows that there is no difference in employment in the first 11 months after hiring between the treatment and the control groups. The point estimate is positive at 0.27 months (+2.6%) but not statistically significant at standard levels. Second, we find no effect on wage trajectories. The differences-in-differences point estimate is again non-significant at 71 euros (3.0%) (Column 2). In the Appendix Figure G3, we also show the trends and did estimates on employment and earnings over time and find that the pattern is similar over time between workers hired into the treatment and the control groups. Thus, there is no direct evidence that hiring anonymously would result in less suitable workers. However, we cannot completely rule out economically meaningful effects in either direction for either outcome variable.

However, in Column (3) of Table 3 Panel D, we show that hiring in the treatment group during the pilot was slower than in the standard hiring processes. The point estimate suggests that, on average, the time from vacancy posting to job start increased by 25 days. This may reflect two technical limitations in the hiring process during the pilot. First, in the anonymized processes, hiring managers could not start interviewing applicants before the application deadline had passed as that would have revealed their identities. Second, the HR had to manually prepare the anonymized applicant summaries to hiring managers which may have prolonged the process. Unfortunately, we do not observe the time of interviews to explore this mechanism further.

5.3 Hiring Managers' Selection into Anonymizing

We next study whether self-selection could explain the difference between our results and previous empirical evidence (Behaghel et al. 2015). During the pilot, treatment was determined by job title, not individual managers. In early 2021, after the pilot period ended, the City of Helsinki made anonymization an opt-in policy. We use this policy change to study which managers opt in as well as hiring outcomes relative to the pilot.

First, Table 4 Panel A describes selection into anonymization by managers during

and after the pilot when it is made an opt-in policy. We find that observable manager-level characteristics are not predictive of the choice to anonymize during the pilot stage. Thus, this is consistent with anonymization being mandated and not based on manager level discretion despite imperfect take-up. This result remains if we control for the city's division or broad occupation-level fixed effects. However, this result changes after anonymization becomes voluntary. Managers who opt in are less likely to be women, and more likely to be younger and Finnish speaking (Columns 4–6). Now, we don't have any direct individual-level measure of discriminatory tendency that would tell us whether those who opt in are more or less likely to discriminate, making it hard to predict what the likely direction this selection has on hiring outcomes. However, these results suggest that hiring managers who individually decide to use anonymization are different from those who do not.

Second, to study the effects on hiring decisions, we use the variation over time in recruitment to the same job title with different hiring practices. We estimate the differences in hiring outcomes for hiring processes that were voluntarily done anonymously vs. via the standard process in a two-way fixed effects specification. Our sample consists only of the pre-period when no anonymizations were made and the post-pilot period when anonymization are voluntary. Interestingly, our estimates in Table 4 Panel B Columns (1–3) suggest foreign-named applicants are not more likely to be called to an interview under voluntary anonymization. This seems to suggest that managers who take up anonymization would have been already inviting more of foreign name-applicants in the absence of the anonymization policy. We also see no increases in hires by foreign-named applicants (Column 4). If anything, it seems that there is a penalty for Finnish-named applicants in the hiring stage, but the gap is not statistical significant at standard levels (Column 6). Overall, these estimates need to be interpreted with some caution as there is no clean control group available. Nevertheless the estimates are consistent with the idea that voluntary anonymization are not similarly effective as mandated anonymization.

5.4 Robustness Checks and Specification Tests

We first replicate our results using alternative measures for immigrant background (Appendix Table F6). Non-native first language, which is independent of the name-based measure but highly correlated, and foreign full names, first names only, or last names only. The estimates are qualitatively very similar to the baseline in all cases.

Second, our estimates assume that hiring in other City of Helsinki job openings is a plausible counterfactual for the pilot titles. The results on hiring decisions remain very similar when we drop individual occupation groups or administrative units (Appendix Table F7). The exceptions are the occupation group “Library, youth, sports, and cultural sector tasks” and the division “Culture and Leisure,” where the estimates are smaller and insignificant. However, as many of the anonymized hiring processes fall into these broader categories, we are left with limited variation in the treatment in these samples.

Third, to test for potential spillovers across hiring processes and which could affect the interpretation of our estimates, we split the control group into vacancies close to and distant from the treated titles (Appendix Section C). Outcomes in the close control group are very similar to those in the distant group during the pilot, consistent with no significant spillovers.

Finally, COVID-19 could have affected treated and control titles differently, but we find no evidence of differential exposure. In particular, event studies of incumbent workers show no differential changes in continued employment or earnings around the pilot (Figure G4), and the number of vacancies in 2020, overall and in treated titles, does not differ systematically from other years (Figure G5).

6 Conclusion

Anonymization seems like the most direct policy response to observed hiring discrimination. For most employers, it is a feasible and a low-cost intervention that can be scaled up

and used in hiring across various types of jobs and occupations. Yet, it has not become widely adopted in the labor markets, suggesting that employers or policy makers believe there are either no clear benefits or that it would make hiring less efficient. Our results suggests that there can be measurable benefits in terms of improving the opportunities of immigrant job applicants, but no detectable costs to match quality. In our context, anonymization also attracted more native job applicants, moderating changes in final recruitment outcomes, which are a function of both application and hiring decisions.

Our results also suggest that implementation details of anonymization are important for their effectiveness and not all policies are likely to work as intended. In combination with previous work, our results suggest that manager's discretion in anonymization can backfire if the intention is to help potentially discriminated groups like immigrant job seekers. But when employers limit this discretion, the policy seems to work exactly in the direction one would predict based on a wealth of correspondence studies. In our context, anonymization process did slow down the hiring process due to administrative frictions, but we believe these could be alleviated with different design choices and automation of the anonymization process, as has been done since the pilot finished at the City of Helsinki. Finally, our results are based on a public sector context, and given previous evidence from ban-the-box policies, there might be reasons to believe that the impacts could differ by sector or by the context in which the anonymization is done. Understanding these differences more systematically seems like a potential avenue for future research.

References

- Agan, A. and Starr, S. (2018). Ban the Box, Criminal Records, and Racial Discrimination: A Field Experiment. *The Quarterly Journal of Economics*, 133(1):191–235.
- Altonji, J. G. and Pierret, C. R. (2001). Employer Learning and Statistical Discrimination. *The Quarterly Journal of Economics*, 116(1):313–350.
- Åslund, O. and Skans, O. N. (2012). Do Anonymous Job Application Procedures Level the Playing Field? *ILR Review*, 65(1):82–107.
- Banfi, S. and Villena-Roldán, B. (2019). Do High-Wage Jobs Attract More Applicants? Directed Search Evidence from the Online Labor Market. *Journal of Labor Economics*, 37(3):715–746.
- Behaghel, L., Crépon, B., and Le Barbanchon, T. (2015). Unintended Effects of Anonymous Résumés. *American Economic Journal: Applied Economics*, 7(3):1–27.
- Belot, M., Kircher, P., and Muller, P. (2022). How Wage Announcements Affect Job Search—A Field Experiment. *American Economic Journal: Macroeconomics*, 14(4):1–67.
- Bertrand, M. and Mullainathan, S. (2004). Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. *American Economic Review*, 94(4):991–1013.
- Cahuc, P., Carcillo, S., Minea, A., and Valfort, M.-A. (2019). When Correspondence Studies Fail to Detect Hiring Discrimination. IZA Discussion Papers 12653, Institute of Labor Economics (IZA).
- Card, D., Colella, F., and Lalive, R. (2023). Gender Preferences in Job Vacancies and Workplace Gender Diversity. RF Berlin - CReAM Discussion Paper Series 2306, Rockwool Foundation Berlin (RF Berlin) - Centre for Research and Analysis of Migration (CReAM).

- Carranza, E., Garlick, R., Orkin, K., and Rankin, N. (2022). Job Search and Hiring with Limited Information about Workseekers' Skills. *American Economic Review*, 112(11):3547–3583.
- Chohlas-Wood, A., Nudell, J., Yao, K., Lin, Z. J., Nyarko, J., and Goel, S. (2021). Blind Justice: Algorithmically Masking Race in Charging Decisions. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, AIES '21, pages 35–45, New York, NY, USA. Association for Computing Machinery.
- Craigie, T.-A. (2020). Ban the Box, Convictions, and Public Employment. *Economic Inquiry*, 58(1):425–445.
- Doleac, J. L. and Hansen, B. (2020). The unintended consequences of "ban the box": Statistical discrimination and employment outcomes when criminal histories are hidden. *Journal of Labor Economics*, 38(2):321–374.
- Farber, H. S. and Gibbons, R. (1996). Learning and Wage Dynamics. *The Quarterly Journal of Economics*, 111(4):1007–1047.
- Flory, J. A., Leibbrandt, A., Rott, C., and Stoddard, O. (2021). Increasing Workplace Diversity: Evidence from a Recruiting Experiment at a Fortune 500 Company. *Journal of Human Resources*, 56(1):73–92.
- Goldin, C. and Rouse, C. (2000). Orchestrating Impartiality: The Impact of "Blind" Auditions on Female Musicians. *The American Economic Review*, 90(4):715–741.
- Ibanez, M. and Riener, G. (2018). Sorting through Affirmative Action: Three Field Experiments in Colombia. *Journal of Labor Economics*, 36(2):437–478.
- Jackson, O. and Zhao, B. (2017). The Effect of Changing Employers' Access to Criminal Histories on Ex-Offenders' Labor Market Outcomes: Evidence from the 2010–2012 Massachusetts Cori Reform.

- Kircher, P. (2022). Job Search in the 21st Century. *Journal of the European Economic Association*, 20(6):2317–2352.
- Kline, P. M., Rose, E. K., and Walters, C. R. (2024). A Discrimination Report Card. NBER Working Papers 32313, National Bureau of Economic Research, Inc.
- Kolev, J., Fuentes-Medel, Y., and Murray, F. (2019). Is Blinded Review Enough? How Gendered Outcomes Arise Even Under Anonymous Evaluation. Working Paper 25759, National Bureau of Economic Research.
- Krause, A., Rinne, U., and Zimmermann, K. F. (2012). Anonymous job applications of fresh Ph.D. economists. *Economics Letters*, 117(2):441–444.
- Kuhn, P. and Shen, K. (2013). Gender Discrimination in Job Ads: Evidence from China *. *The Quarterly Journal of Economics*, 128(1):287–336.
- Kuhn, P. and Shen, K. (2023). What Happens When Employers Can No Longer Discriminate in Job Ads? *American Economic Review*, 113(4):1013–1048.
- Kuhn, P., Shen, K., and Zhang, S. (2020). Gender-targeted job ads in the recruitment process: Facts from a chinese job board. *Journal of Development Economics*, 147:102531.
- Lang, K. and Lehmann, J.-Y. K. (2012). Racial Discrimination in the Labor Market: Theory and Empirics. *Journal of Economic Literature*, 50(4):959–1006.
- Leibbrandt, A. and List, J. (2018). Do Equal Employment Opportunity Statements Backfire? Evidence From A Natural Field Experiment On Job-Entry Decisions. Technical Report w25035, National Bureau of Economic Research, Cambridge, MA.
- Neumark, D. (2018). Experimental Research on Labor Market Discrimination. *Journal of Economic Literature*, 56(3):799–866.
- Neumark, D. (2024). Age Discrimination in Hiring: Evidence from Age-Blind versus Non-Age-Blind Hiring Procedures. *Journal of Human Resources*, 59(1):1–34.

- Pallais, A. (2014). Inefficient Hiring in Entry-Level Labor Markets. *American Economic Review*, 104(11):3565–3599.
- Pyle, B. (2024). A Second Look: Local Labor Markets and the Impact of Ban the Box Policies after Criminal Legal Involvement.
- Riach, P. A. and Rich, J. (2002). Field Experiments of Discrimination in the Market Place. *Economic Journal*, 112(483):480–518.
- Rich, J. (2014). What Do Field Experiments of Discrimination in Markets Tell Us? A Meta Analysis of Studies Conducted since 2000. IZA Discussion Papers 8584, Institute of Labor Economics (IZA).
- Rose, E. K. (2021). Does Banning the Box Help Ex-Offenders Get Jobs? Evaluating the Effects of a Prominent Example. *Journal of Labor Economics*, 39(1):79–113.
- Shoag, D. and Veuger, S. (2021). Ban-the-box measures help high-crime neighborhoods. *The Journal of Law and Economics*, 64(1):85–105.
- Zschirnt, E. and Ruedin, D. (2016). Ethnic discrimination in hiring decisions: a meta-analysis of correspondence tests 1990–2015. *Journal of Ethnic and Migration Studies*, 42(7):1115–1134.

Tables

Table 1: Pre-Treatment Characteristics

	All (1)	Control (2)	Treated (3)	$\Delta T - \Delta C$ (4)	S.E. (5)
Panel A: Job Characteristics, $N = 2,489$					
No. Applicants	10.81	10.64	18.98	-0.13	(3.43)
Temporary Contract	0.41	0.41	0.41	0.32**	(0.15)
Panel B: Applicant Characteristics, $N = 25,940$					
Woman	0.78	0.79	0.75	-0.04	(0.03)
Foreign First Name	0.07	0.07	0.12	-0.00	(0.02)
Foreign Last Name	0.16	0.16	0.25	0.04*	(0.02)
Any Foreign Name	0.19	0.19	0.29	0.03	(0.03)
Finnish Native Language	0.92	0.92	0.87	0.01	(0.02)
Age	38.32	38.30	38.75	0.67	(0.66)
Bachelor's Degree	0.82	0.83	0.50	-0.11	(0.08)
Experience (Years)	10.09	10.11	9.40	0.49	(0.57)
Interview	0.09	0.09	0.08	-0.02	(0.02)
Helsinki Worker	0.46	0.46	0.40	-0.05	(0.04)
Panel C: Recruit Characteristics, $N = 1,393$					
Woman	0.80	0.81	0.70	-0.01	(0.17)
Foreign First Name	0.07	0.08	0.03	-0.11	(0.07)
Foreign Last Name	0.15	0.15	0.17	0.09	(0.14)
Any Foreign Name	0.18	0.18	0.20	-0.00	(0.14)
Finnish Native Language	0.93	0.93	0.97	0.10	(0.07)
Age	38.53	38.60	35.41	-2.57	(2.17)
Bachelor's Degree	0.76	0.76	0.70	-0.46**	(0.19)
Experience (Years)	10.11	10.13	9.58	-2.70	(1.91)
Helsinki Worker	0.64	0.64	0.68	-0.21	(0.16)

Notes: This table shows the characteristics of job openings, applicants and recruits to job openings at the City of Helsinki before the hiring pilot (September 2019 to January 2020). Column (1) shows the average characteristics for all job openings, and Columns (2) and (3) show the characteristics separately for job openings in the control and treated groups. Column (4) shows the pre-treatment difference in differences between the treated (ΔT) and the control (ΔC) groups. The pre-period has been divided into two 10-week time periods. The associated standard errors are in Column (5). * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table 2: Difference-in-Difference Estimates on Hiring and Calls to Interviews

	P(Interview)							P(Hired)						
	Foreign		Finnish		Pooled			Foreign		Finnish		Pooled		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Panel A: Immigrant background														
Treat X Post	0.045** (0.022)	0.045** (0.022)	-0.007 (0.015)	-0.009 (0.015)	-0.006 (0.015)	-0.007 (0.015)		0.031** (0.014)	0.031** (0.014)	-0.024*** (0.008)	-0.024*** (0.009)	-0.023*** (0.008)	-0.023*** (0.008)	
Foreign X Treat X Post					0.052** (0.021)	0.052** (0.021)	0.050** (0.020)					0.050*** (0.016)	0.049*** (0.015)	0.045*** (0.016)
Foreign X Post					-0.015*** (0.006)	-0.015*** (0.006)	-0.010 (0.006)					-0.010** (0.005)	-0.010** (0.005)	-0.007 (0.005)
Foreign X Treat					-0.025 (0.016)	-0.026* (0.016)	-0.019 (0.016)					-0.028** (0.013)	-0.029** (0.013)	-0.023* (0.014)
N	16,593	16,593	58,102	58,102	74,802	74,802	73,877	16,593	16,593	58,102	58,102	74,802	74,802	73,877
Foreign Mean	0.063	0.063			0.063	0.063	0.063	0.031	0.031			0.031	0.031	0.031
Finnish Mean			0.086	0.086	0.086	0.086	0.086			0.051	0.051	0.051	0.051	0.051
	P(Interview)							P(Hired)						
	Women		Men		Pooled			Women		Men		Pooled		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Panel B: Gender														
Treat X Post	0.006 (0.016)	0.005 (0.017)	0.028 (0.030)	0.026 (0.030)	0.026 (0.028)	0.026 (0.028)		-0.006 (0.007)	-0.007 (0.007)	-0.016 (0.018)	-0.017 (0.018)	-0.009 (0.019)	-0.009 (0.019)	
Woman X Treat X Post					-0.023 (0.033)	-0.024 (0.032)	-0.033 (0.030)					0.001 (0.022)	-0.000 (0.022)	0.001 (0.024)
Woman X Post					0.010 (0.006)	0.009 (0.006)	0.019*** (0.006)					0.013*** (0.004)	0.012*** (0.004)	0.015*** (0.005)
Woman X Treat					-0.007 (0.026)	-0.003 (0.025)	0.002 (0.024)					-0.017 (0.021)	-0.013 (0.021)	-0.021 (0.023)
N	58,049	58,049	16,663	16,663	74,802	74,802	73,877	58,049	58,049	16,663	16,663	74,802	74,802	73,877
Women Mean	0.086	0.086			0.086	0.086	0.086	0.042	0.042			0.042	0.042	0.042
Men Mean			0.059	0.059	0.059	0.059	0.059			0.054	0.054	0.054	0.054	0.054
Job FE	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	
Vacancy FE							✓							✓
Demographic controls		✓		✓		✓			✓		✓		✓	
Skill controls		✓		✓		✓			✓		✓		✓	

Notes: This table shows the difference-in-difference estimates using the specifications in Equation 1 and analogous triple difference specification testing for the gap by applicant groups. Treat $\times$ Post is the estimated difference-in-differences coefficient, an intention-to-treat estimate of assignment to the anonymization pilot. In the subgroup columns (1–4 and 8–11) the sample is restricted to applicants from that group and Treat $\times$ Post gives the within-group treatment effect. In the pooled columns (5–7 and 12–14) all applicants are included and the triple interaction Foreign $\times$ Post $\times$ Treat (Panel A) or Woman $\times$ Post $\times$ Treat (Panel B) tests whether the treatment effect differs between the two groups. Specifications with "Job FE" absorb job-title fixed effects; the last column of each outcome block uses vacancy fixed effects instead. Demographic controls are gender or name origin (whichever is not the panel's subgroup variable), age, and age squared. Skill controls are work experience, experience squared, and dummies for highest completed education level. Foreign/Women Mean and Finnish/Men Mean are pre-treatment means of the outcome among applicants from each subgroup in treated vacancies; in the pooled columns both are reported. Subgroup regressions drop job titles with a single applicant from that group. Robust standard errors in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table 3: Difference-in-Difference Estimates on Applications and Post-Hiring Outcomes

Panel A: Applications and the Hiring Process					
	(1) No. Applicants	(2) No. Finnish	(3) No. Foreign	(4) No. Interviewed	(5) P(Hire)
Treat X Post	5.692** (2.665)	4.808** (1.992)	0.884 (0.812)	0.425 (0.344)	0.030 (0.033)
<i>N</i>	7,512	7,512	7,512	7,512	7,512
Outcome Mean	18.980	12.796	6.184	1.449	0.837
Panel B: Applicant Characteristics					
	(1) P(Foreign)	(2) P(Woman)	(3) P(High Edu)	(4) Work Experience	(5) Age
Treat X Post	-0.052*** (0.018)	0.055*** (0.021)	0.090** (0.038)	-0.229 (0.552)	-1.552*** (0.535)
<i>N</i>	7,512	7,442	5,624	7,512	6,863
Outcome Mean	0.332	0.743	0.633	7.056	36.481
Panel C: Hiree Characteristics					
	(1) P(Foreign)	(2) P(Woman)	(3) P(High Edu)	(4) Work Experience	(5) Age
Treat X Post	0.040 (0.095)	0.111 (0.099)	-0.036 (0.117)	-1.964* (1.171)	-1.378 (1.964)
<i>N</i>	3,861	3,713	2,933	3,861	3,516
Outcome Mean	0.200	0.700	0.704	9.031	35.415
Panel D: Post-Hiring Outcomes					
	(1) Months Employed	(2) Monthly Wage	(3) Days to Job Start		
Treat X Post	0.265 (0.451)	70.567 (67.948)	25.335*** (6.165)		
<i>N</i>	3,214	3,214	3,214		
Outcome Mean	10.077	2,396	48.538		

Notes: Panel A shows difference-in-differences estimates on job applications aggregated to the job opening level. Panel B shows the difference-in-differences estimates on the applicant characteristics at the job level. Panels A and B are estimated at the vacancy level. Columns 1 and 2 are weighted by the vacancy's number of applicants so that the estimates reflect the average applicant's experience rather than the average vacancy. Panel C shows the difference-in-differences estimates on the hiree characteristics at the job level. Panel C is also estimated at the vacancy level but unweighted, since the outcomes already represent vacancy-level averages over the (typically small) number of hires. Panel D shows the difference-in-differences estimates on post-hiring outcomes for the hired applicants. Months employed in Column (1) and average monthly wage in Column (2) are measured in the first 11 months of employment. Column (3) is the time from the start of the application period to the first salary month. Panel D is estimated at the individual level. The outcome means are the pre-treatment period means for the treated group. Robust standard errors in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

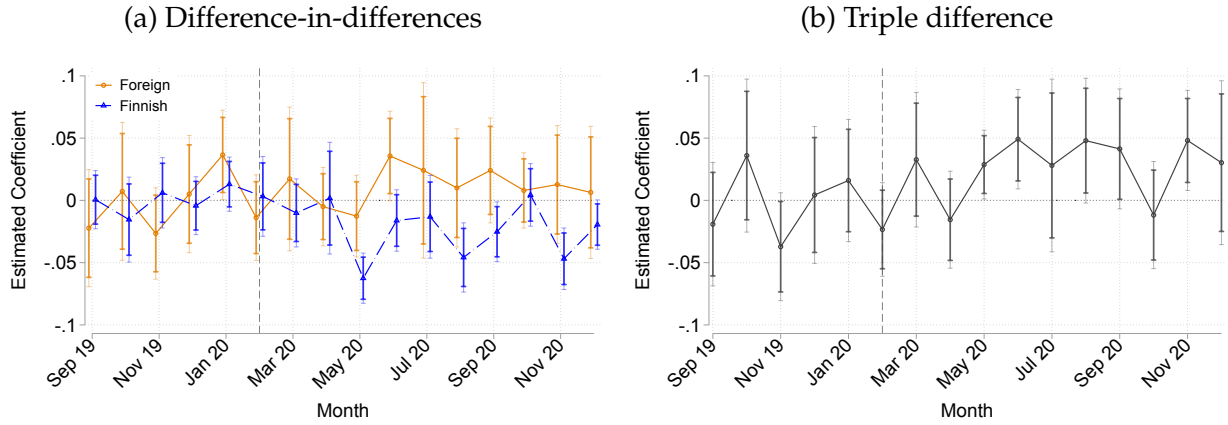
Table 4: Voluntary Anonymization: Managers' Self-Selection and Hiring Outcomes

Panel A: Hiring managers' self-selection into anonymization						
	Pilot (mandated)			Post-pilot (voluntary)		
	(1)	(2)	(3)	(4)	(5)	(6)
Woman recruiter	-0.133 (0.177)	-0.013 (0.199)	0.087 (0.180)	-1.793** (0.849)	-1.143 (0.723)	-1.195* (0.712)
Recruiter's wage (1000)	-0.181 (0.129)	-0.202 (0.137)	-0.283* (0.159)	0.260 (0.177)	0.135 (0.125)	0.018 (0.111)
Recruiter's age	0.022 (0.020)	0.022 (0.019)	0.025 (0.019)	-0.067*** (0.025)	-0.046*** (0.017)	-0.047*** (0.015)
Native language Finnish	0.368* (0.219)	0.297 (0.206)	0.351 (0.290)	1.536** (0.652)	0.691 (0.521)	1.137* (0.638)
Outcome mean	0.417	0.417	0.417	2.076	2.076	2.076
N	6,230	6,230	6,230	22,163	22,163	22,163
Division FE		✓	✓		✓	✓
Occupation FE			✓			✓
Panel B: Hiring outcomes under voluntary anonymization						
	P(Interview)			P(Hired)		
	Foreign (1)	Finnish (2)	Pooled (3)	Foreign (4)	Finnish (5)	Pooled (6)
Anonymous	-0.003 (0.037)	-0.011 (0.022)	-0.009 (0.021)	-0.016 (0.030)	-0.057*** (0.018)	-0.050*** (0.017)
Anonymous × Foreign			0.008 (0.032)			0.029 (0.026)
Foreign			-0.018*** (0.003)			-0.013*** (0.003)
N	9,980	36,976	47,043	9,980	36,976	47,043
Demographic controls	✓	✓	✓	✓	✓	✓
Job FE	✓	✓	✓	✓	✓	✓
Month FE	✓	✓	✓	✓	✓	✓

Notes: Panel A reports linear probability estimates of whether a recruitment process was anonymized, expressed in percentage points (0–100), as a function of recruiter characteristics. Columns (1)–(3) cover the pilot period, when anonymization was mandated by job title and not subject to individual managers' discretion; Columns (4)–(6) cover the post-pilot period, when anonymization was voluntary. All columns control for the number of vacancies and whether the contract is temporary; the final month of observation for vacancies is August 2023. Panel B reports estimates of hiring outcomes in voluntarily anonymized relative to standard recruitment processes from the two-way fixed effects specification $Y_{ijt} = \alpha + \beta \text{Anonymous}_j + \lambda_t + \theta_j + \varepsilon_{ijt}$, where Anonymous_j indicates that the process was voluntarily anonymized; Columns (3) and (6) add an interaction with immigrant background. The sample excludes recruitment processes in the mandated anonymization period. Robust standard errors in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Figures

Figure 1: Event-Study Estimates on Hiring Decisions



Notes: These figures show the event-study estimates on hiring decisions. Panel A shows the estimates separately for applicants with foreign (yellow) and Finnish (blue) names from the following specification: $Y_{ijt} = \alpha + \sum_{k=-5}^{11} \beta_k \text{Treated}_j \times \mathbf{1}\{t - \text{Feb2020} = k\} + \lambda_t + \theta_j + \varepsilon_{ijt}$, where we restrict $\sum_{k=-5}^{-1} \beta_k = 0$. Panel B shows analogous estimates for the gap between applicants with foreign relative to applicants with Finnish names from a triple-diff specification. Figures show 90 and 95 percent confidence intervals around point estimates.

Supplementary Material

A Institutional setting and data

The City of Helsinki provides an interesting context for the anonymization pilot for at least three reasons. First, as Finland's largest employer, with over 37,000 employees, there is significant variety in the types of sectors and occupations that people need to be hired into. Second, the immigrant population in Helsinki has grown significantly from 2 percent in 1990 to 19 percent in 2022³, increasing returns to reducing hiring frictions for immigrants. Third, access to administrative data allows us to follow the whole recruitment process, from the job opening to job applications and hiring decisions, as well as post-hiring outcomes.

The pilot was implemented in two stages. First, the initial plan was to randomize anonymization among job openings that belonged to the selected job titles. However, due to a relatively low number of job openings to those job titles in the beginning of the pilot, the city decided to stop randomizing and subjected all job openings in the selected job titles to an anonymous hiring process starting in May 2020. The main justification was to learn from a larger number of anonymized hiring processes at a practical level. As the initial randomization stage was very short and includes an insufficient number of observations for a separate analysis, we pool together all job openings in the treated job titles, whether or not they were assigned to anonymization or not.

Take-up of anonymization was not perfect during the pilot. The take-up rate was lower in the first months of the pilot when only randomly selected job openings were intended for anonymization (see Figure G1). Overall, 44 of 99 or 44 percent of the recruitment processes in the treatment group were conducted anonymously. After discussing with the city HR department, we expect that part of the imperfect take-up is driven by a lack of resources in the departments responsible for hiring specific job titles. In particular, anonymization required extra work by the HR department. However, we cannot exclude the possibility that some of the imperfect take-up might be driven by hiring managers' decisions, an issue we discuss in Section 5.3.

In our empirical analysis, we focus on job openings with a single vacancy. This covers more than 85 percent of all job openings and allows us to exclude job openings that could include multiple job titles, thereby reducing measurement error in whether hiring processes belong to the treatment or control groups. In the individual-level analysis, we also use data on applicants whose complete background information (date of birth, education,

³Official Statistics of Finland, OSF: Population structure, Statistics Finland Referenced: 19.2.2024.

and job experience) we observe. This ensures a consistent sample across analyses with or without individual-level controls. Additionally, we require birth dates to match applicants to payroll data, maintaining consistency in these separate analyses. Importantly, our estimates of the main hiring outcomes remain qualitatively unchanged whether or not we restrict the sample to these individuals.

Our data comes from two separate sources: the recruiting system and the payroll data. There is a mismatch between two sources in our data on whether someone was hired. In some cases, we do not observe an individual being hired even if the job opening was closed and tagged as "someone hired". We expect most of the mismatch to result from missing information at the individual level, not from wrong tagging at the job opening level. In fact, we observe that during the pilot, this mismatch is larger in the treatment group than in the control group. The Diff-in-Diff estimate on having data on the hiree is -0.235 with a baseline of 0.861. This is a non-negligible share of all hirees. We expect this to be mostly driven by the fact that during the pilot phase, the anonymization and the calls to interviews were done in practice by HR, and the hiring managers did not always use the recruitment system to keep track of the final hiring outcomes at the individual level (see Section 2). This issue does not concern our data on calls for interviews, which are recorded in a credible manner, as discussed in Section 2.

We define the job start month using the first observed positive monthly earnings after the job vacancy has been posted. For workers who already had a work contract with the City of Helsinki prior to hiring, we define the job start using the first earnings from a new job spell observed in the payroll data.

For the wage data and employment length to be informative about match quality between the workers and jobs, there should be potential individual level variation. First, while collective agreements cover a large share of the Finnish work contracts and set common wage floors to workers in the same job, there is still significant variation in wages and earnings conditional on the job title. The standard deviation for starting wages in our data is €1,235 and for wages net of the average wage in the job title €833. The job characteristics directly explain only one third of the variation in starting wages. Our data also includes any promotions and wage increases during the follow-up period. Second, while temporary contracts are common (over 40% in our data), all contracts have a probation period which is 6 months or one half of the contract length if the contract is less than 12 months. We also follow individual workers if they sign a new contract extension. This means that even for workers who signed temporary or fixed term contracts, there is room for individual level variation depending on the match quality.

B Linking Hires and Hiring Managers to Payroll Data

To measure the earnings and employment duration of new hires, we match all hires to an administrative payroll data set containing all City of Helsinki employees and their wages each month. This data set contains information on the base wage corresponding to their job title, the total wage (including overtime, seniority, and bonus payments), and monthly earnings (actual monthly payment, which may depend on hours worked). We measure employment by months in which an employee had positive earnings from the city of Helsinki. The payroll data is monthly from September 2019 to February 2022. For the linkage, we use vacancies, whose application period started between September 2019 and December 2020.

To link these two data, we use the information on the names and dates of birth of hires. The names do not always match precisely across the datasets. Applicants may spell their name in a different way or choose to use only part of their full name when sending their job applications. For the linking process, we separate the first names and last names in the applicant data and keep the full name in the payroll data. We transform the names into LATIN-ASCII characters to remove any special characters. We use fuzzy record linkage to merge the two datasets. The results of the linkage are presented in Table B1. There are 3,568 hired applicants who have reported their date of birth in the application. There are 4,176 hire events—an individual may have been hired multiple times. After linking hires, we have 3,580 matches in total.

For the hiring managers, we observe the name in the hiring platform and use the *reclink* package in Stata to match them to the payroll data to get their characteristics. The payroll data gives us the hiring managers' wage, age, and native language. We are able to match 87% managers to a worker in the payroll data.

C Spillover Analysis

A potential threat to the interpretation of our difference-in-differences estimates is that the pilot could create spillovers across hiring processes. For example, if anonymization attracts some applicants who would have otherwise applied to jobs in the control group, then the estimated treatment effects on hiring outcomes could partly reflect a reallocation of applicants across vacancies. This would violate the SUTVA assumption and potentially bias our estimates if it affected the composition or outcomes of the control group, contaminating the counterfactual.

Now, we expect that if these spillovers exist, they should be concentrated among control vacancies that are close substitutes for the treated vacancies. In particular, applicants

Table B1: Linking the Application Data to the Payroll Data

Data	N	Share
Employees	76,722	
Hired applicants	4,005	
reporting both a name and a date of birth	3,568	
Hires linked to an employee	3,444	86.0%
Hire events (hired applications)	4,176	
Hire events dated to a hire month	3,580	85.7%
Hire events in the analysis sample	3,214	
Recruiters	2,765	
Recruiters matched	2,413	87.3%

Notes: The table reports the number of employees in the payroll data, applicants, and the match rate. Hire events are applicant–posting pairs in which the applicant was hired; an applicant may have been hired for more than one posting, so hire events exceed hired applicants. The match rate in the upper panel is taken over all hired applicants, including those who report no usable date of birth and so cannot be compared to the payroll at any threshold; among the 3,568 who report both a name and a date of birth it is 96.5%. The analysis sample is dated hire events whose application period opened between September 2019 and December 2020; the eleven-month outcome window ends within the payroll extract for all of them.

are more likely to apply to occupations that are close to one another in terms of required skills, tasks, and qualifications than between occupations that are more distant. To make some progress, we split the control group into two groups based on their occupational proximity to the treated job titles. We use the City of Helsinki’s occupational classification to assign every job title to a broad occupational group. *Close controls* are control job titles that fall in the same broad occupational group as at least one treated job title and *distant controls* are control job titles in occupational groups that contain no treated job titles. If anonymization there were important spillovers, the close control group should be more exposed and the distant control group should be less exposed.

Table F4 shows the resulting split of the control group across occupational groups, together with the number of applicants and vacancies in each cell. Treated job titles appear only in some occupational groups (for example, customer service and office tasks, expert roles, supervisory and managerial tasks, ICT tasks, and library, youth, sports, and cultural sector tasks), and it is the control job titles within these same groups that we classify as close controls. Out of 88,651 applicants and 8,012 vacancies in the control group, 31,500 applicants and 1,380 vacancies fall in the close control group, with the remainder classified as distant controls.

To test for spillovers, we re-estimate our main specification while allowing the treat-

ment effect to differ between close and distant controls. Concretely, we estimate

$$Y_{ijt} = \alpha + \beta_1 \text{Treated}_j \times \text{Post}_t + \beta_2 \text{Close}_j \times \text{Post}_t + \delta X_{ijt} + \lambda_t + \theta_j + \varepsilon_{ijt}, \quad (2)$$

where Close_j is a dummy equal to one for close control job titles and zero for distant control job titles, and all other terms are defined as in Equation 1. The coefficient β_1 recovers the intention-to-treat estimate of assignment on the treated openings, identified relative to the distant controls. The coefficient of interest for the spillover test is β_2 , which measures whether hiring outcomes in the close control group differ from those in the distant control group during the pilot. Under the null of no spillovers, anonymization of the treated openings should leave outcomes in the close control group unchanged relative to the distant controls, so that $\beta_2 \approx 0$. A non-zero β_2 would instead indicate that outcomes in occupations closer to the treated ones were possibly affected by spillovers.

Table F5 shows the estimates separately by applicant group and, in the pooled columns, include the analogous triple-difference interaction that tests whether the treatment effect differs by applicant background, exactly as in our main specification (Equation 1). Hiring outcomes in the close control group are very similar to those in the distant control group during the pilot, so that the estimated spillover term is small and statistically insignificant. We read this as evidence that the pilot did not meaningfully reallocate applications across nearby hiring processes, and that our main estimates are not driven by contamination of the control group.

D Testing Disclosure of Group Identity in Anonymized Hiring

The anonymization pilot removed identifying information from the structured fields of each application before the file was forwarded to the hiring manager (see Section 2). However, applications also contain open text fields, most importantly the motivation letter, that were forwarded to managers and that were not screened for anonymization. A natural concern is that applicants might reveal identifying information in these open text fields (name, gender, national or ethnic background, or the educational institution they attended) and thereby undermine the anonymization treatment. If such disclosure were common, the effective intensity of the treatment would be weaker than intended, and our estimates would understate the effect of fully concealing applicant identity.

To gauge how often identifying information might be disclosed in the open text fields, we did a manual audit. In particular, we drew a random sample of 100 applications submitted in 2019–2020, split evenly between 50 applications to treated (anonymized) recruitments and 50 applications to control recruitments. Because our primary interest is

in the disclosure of immigrant background, we ensure that each group contains applicants classified as having a foreign background, with 12 such applicants in each of the two groups. We then manually read the open text fields in each application and classify whether the text reveals, or makes it possible to infer, identifying information. Specifically, we record whether the text discloses (i) the applicant’s name, (ii) the applicant’s gender, and (iii) other potentially identifying details such as the educational institution attended. For background in particular, we distinguish between text that clearly identifies a foreign background and text that makes a foreign background plausible but not necessarily certain.

Table D1 summarizes the results and suggests that disclosure of identifying information in the open text fields is rare. In the treated group, no applicant clearly revealed a foreign background and none was coded as possibly revealing a foreign background, while gender could be inferred from the text for a small number of applicants. The pattern is similar in the control group. Overall, the audit indicates that very few applicants disclose identifying information in free-text responses. We therefore conclude that open text fields are unlikely to have materially weakened the anonymization treatment, and that the structured-field anonymization performed by the HR department was a binding constraint on the information available to hiring managers at the screening stage.

Table D1: Disclosure of Identifying Information in Open Text Fields

Revealed information	Treatment	Control
Foreign background identifiable	0	1
Possibly foreign background	0	1
Gender identifiable (female)	1	1
Gender identifiable (male)	1	0

Notes: The table reports the number of applications, out of a manual audit of 100 applications submitted in 2019–2020 (50 to treated recruitments and 50 to control recruitments, each including 12 applicants classified as having a foreign background), in which the open text fields reveal or make it possible to infer the indicated identifying information. “Foreign background identifiable” counts applications whose text clearly reveals a foreign background; “Possibly foreign background” counts applications whose text makes a foreign background plausible but not certain.

E Hiring Conditional on Reaching the Interview Stage

Our main specification estimates the effect of anonymization on hiring outcomes among all applicants. This is the policy-relevant object of interest, since it captures the combined

effect of anonymization operating through both the call-to-interview margin and the final hiring decision. An influential strand of the discrimination literature, however, focuses on hiring *conditional on reaching the interview stage* (e.g., Neumark 2018, 2024). This alternative estimand isolates the hiring decision among candidates who have already been interviewed, abstracting from the screening stage at which names and other identifying information are first used. This appendix re-estimates our main effects following this approach, both to assess the robustness of our conclusions and to make our results directly comparable to this prior work.

Concretely, we re-estimate the difference-in-differences and triple-difference specifications from Equation 1, but restrict the sample used for the hiring outcome to applicants who reached the interview stage. The call-to-interview regressions are unchanged. We report estimates for applicants with a foreign name, exactly as in our main analysis, and we additionally study older applicants, whom we define as those aged 50 or above. We use this threshold because it captures the margin at which age discrimination is typically most salient, while higher thresholds (e.g., 55 or above) sharply reduce the sample of treated applicants and the associated statistical power.

Table E1 reports the conditional-on-interview estimates for foreign-named applicants (Panel A). Panel B reports the difference-in-differences and triple-difference specifications from Equation 1 for workers over and under 50. Panel C reports the corresponding estimates by age, restricting the hire sample to those who were interviewed.

Two patterns stand out. First, for foreign-named applicants the positive effect on calls to interviews carries over to hiring decisions, and conditioning on the interview stage does not overturn this conclusion. Second, older applicants receive more calls to interviews under anonymization, but this does not translate into more hires conditional on being interviewed; the conditional hiring estimates are negative. This pattern is qualitatively consistent with the findings of Neumark (2024), and suggests that the mechanisms behind age and immigrant-background discrimination may differ, despite the differences between our public-sector setting and the private-sector settings studied in much of the prior literature. This is also broadly consistent with the modest, if imprecisely estimated, reduction in hires' average age reported in Section 5.1 (Table 3, Panel C), suggesting some role for age alongside origin as a factor affected by anonymization.

These estimates should be interpreted with some caution. Conditioning on reaching the interview stage selects on a post-treatment outcome, so the conditional-on-interview sample need not be balanced across the treatment and control groups even if the underlying assignment is exogenous. For this reason we treat the conditional estimates as a complement to, rather than a substitute for, our main unconditional estimates. The broad

consistency of the two approaches nonetheless reinforces the robustness of our main conclusions.

Table E1: Foreign Hiring and Hiring by Age Baseline and Conditional-on-Interview Specification

	P(Interview)							P(Hired)						
	Foreign		Finnish		Pooled			Foreign		Finnish		Pooled		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Panel A: Immigrant background, hiring conditional on reaching the interview stage														
Treat X Post	0.045** (0.022)	0.045** (0.022)	-0.007 (0.015)	-0.009 (0.015)	-0.006 (0.015)	-0.007 (0.015)		0.269* (0.139)	0.272** (0.125)	-0.182** (0.090)	-0.189** (0.091)	-0.196** (0.089)	-0.201** (0.089)	
Foreign X Treat X Post					0.052** (0.021)	0.052** (0.021)	0.050** (0.020)					0.240 (0.189)	0.244 (0.184)	0.105 (0.198)
N	16,593	16,593	58,102	58,102	74,802	74,802	73,877	1,073	1,073	5,216	5,216	6,394	6,394	3,385
Foreign Mean	0.063	0.063			0.063	0.063	0.063	0.500	0.500			0.500	0.500	0.500
Finnish Mean			0.086	0.086	0.086	0.086	0.086			0.593	0.593	0.593	0.593	0.593
	P(Interview)							P(Hired)						
	≥ 50		< 50		Pooled			≥ 50		< 50		Pooled		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Panel B: Age, baseline specification														
Treat X Post	0.040** (0.017)	0.035** (0.017)	0.001 (0.016)	0.001 (0.016)	0.002 (0.016)	0.001 (0.016)		0.002 (0.011)	-0.000 (0.011)	-0.011 (0.007)	-0.010 (0.007)	-0.011* (0.007)	-0.011 (0.007)	
Over 50 X Treat X Post					0.033 (0.023)	0.027 (0.023)	0.027 (0.022)					0.009 (0.015)	0.005 (0.015)	-0.002 (0.014)
N	15,106	15,106	59,590	59,590	74,802	74,802	73,877	15,106	15,106	59,590	59,590	74,802	74,802	73,877
≥ 50 Mean	0.036	0.036			0.036	0.036	0.036	0.018	0.018			0.018	0.018	0.018
< 50 Mean			0.081	0.081	0.081	0.081	0.081			0.046	0.046	0.046	0.046	0.046
	P(Interview)							P(Hired)						
	≥ 50		< 50		Pooled			≥ 50		< 50		Pooled		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Panel C: Age, hiring conditional on reaching the interview stage														
Treat X Post	0.040** (0.017)	0.035** (0.017)	0.001 (0.016)	0.001 (0.016)	0.002 (0.016)	0.001 (0.016)		-0.414* (0.235)	-0.407* (0.224)	-0.089 (0.094)	-0.096 (0.094)	-0.103 (0.094)	-0.111 (0.094)	
Over 50 X Treat X Post					0.033 (0.023)	0.027 (0.023)	0.027 (0.022)					-0.265 (0.168)	-0.278* (0.168)	-0.413** (0.194)
N	15,106	15,106	59,590	59,590	74,802	74,802	73,877	1,189	1,189	5,103	5,103	6,394	6,394	3,385
≥ 50 Mean	0.036	0.036			0.036	0.036	0.036	0.500	0.500			0.500	0.500	0.500
< 50 Mean			0.081	0.081	0.081	0.081	0.081			0.565	0.565	0.565	0.565	0.565
Job FE	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	
Vacancy FE							✓							✓
Demographic controls		✓		✓		✓			✓		✓		✓	
Skill controls		✓		✓		✓			✓		✓		✓	

Notes: Panel A re-estimates the difference-in-differences and triple-difference specifications of Equation 1 with the hiring outcome restricted to applicants who reached the interview stage. Panel A reports estimates by name origin (foreign vs. Finnish). The dependent variable in Columns (1)–(7) is the probability of being called to an interview and in Columns (8)–(14) the probability of being hired conditional on having been interviewed. In the subgroup columns the sample is restricted to applicants from that group and Treat X Post is the within-group treatment effect; in the pooled columns all applicants are included and the triple interaction (Foreign X Treat X Post in Panel A) tests whether the effect differs across groups. Specifications with “Job FE” absorb job-title fixed effects; the last column of each outcome block uses vacancy fixed effects instead. Demographic controls are the complementary group indicator, age, and age squared; skill controls are work experience, experience squared, and dummies for highest completed education. Panel B reports difference-in-differences and triple-difference estimates of the effect of anonymization on older applicants, defined as those aged 50 or above. Panel B uses our baseline specification, in which the hiring outcome is defined over all applicants; Panel C restricts the hiring outcome to applicants who reached the interview stage, following the conditional-on-interview approach of Neumark (2024). In the subgroup columns the sample is restricted to applicants in the indicated age group and Treat X Post is the within-group treatment effect; in the pooled columns all applicants are included and the triple interaction Over 50 X Treat X Post tests whether the effect differs by age group. Group means are pre-treatment means of the outcome among applicants in treated vacancies. Robust standard errors in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

F Tables

Table F1: Number of Observations by Application Start Month

Month	Applicant Level			Vacancy Level		
	Control	Treated	Total	Control	Treated	Total
2019m9	2,260	132	2,392	263	9	272
2019m10	4,202	167	4,369	463	10	473
2019m11	3,829	257	4,086	414	14	428
2019m12	3,024	132	3,156	415	6	421
2020m1	12,656	242	12,898	885	10	895
2020m2	5,812	111	5,923	485	6	491
2020m3	4,108	162	4,270	381	6	387
2020m4	13,148	144	13,292	798	5	803
2020m5	9,288	34	9,322	660	1	661
2020m6	5,661	533	6,194	461	22	483
2020m7	4,008	215	4,223	386	9	395
2020m8	4,749	193	4,942	451	7	458
2020m9	4,248	390	4,638	455	14	469
2020m10	4,285	276	4,561	451	10	461
2020m11	4,737	548	5,285	521	14	535
2020m12	3,524	167	3,691	445	5	450
Total	89,539	3,703	93,242	7,934	148	8,082

Table F2: Main Results Excluding Randomization Phase (02/2020-05/2020)

	P(Interview)							P(Hired)						
	Foreign		Finnish		Pooled			Foreign		Finnish		Pooled		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Panel A: Foreign														
Treat X Post	0.061** (0.024)	0.060** (0.024)	-0.016 (0.016)	-0.018 (0.016)	-0.016 (0.015)	-0.018 (0.016)		0.038*** (0.015)	0.039*** (0.015)	-0.030*** (0.008)	-0.030*** (0.009)	-0.029*** (0.008)	-0.030*** (0.008)	
Foreign X Treat X Post					0.081*** (0.023)	0.082*** (0.022)	0.071*** (0.022)					0.064*** (0.017)	0.063*** (0.016)	0.054*** (0.017)
N	10,939	10,939	36,452	36,452	47,488	47,488	46,776	10,939	10,939	36,452	36,452	47,488	47,488	46,776
Foreign Mean	0.063	0.063			0.063	0.063	0.063	0.031	0.031			0.031	0.031	0.031
Finnish Mean			0.086	0.086	0.086	0.086	0.086			0.051	0.051	0.051	0.051	0.051
Conditional Mean					-0.021	-0.034	-0.022					-0.020	-0.032	-0.019
	P(Interview)							P(Hired)						
	Women		Men		Pooled			Women		Men		Pooled		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Panel B: Women														
Treat X Post	0.003 (0.017)	0.000 (0.017)	0.032 (0.033)	0.030 (0.033)	0.028 (0.030)	0.027 (0.030)		-0.009 (0.007)	-0.011 (0.008)	-0.016 (0.019)	-0.017 (0.020)	-0.007 (0.019)	-0.007 (0.020)	
Woman X Treat X Post					-0.026 (0.035)	-0.028 (0.034)	-0.038 (0.031)					-0.005 (0.023)	-0.007 (0.023)	-0.004 (0.025)
N	37,242	37,242	10,172	10,172	47,488	47,488	46,776	37,242	37,242	10,172	10,172	47,488	47,488	46,776
Women Mean	0.086	0.086			0.086	0.086	0.086	0.042	0.042			0.042	0.042	0.042
Men Mean			0.059	0.059	0.059	0.059	0.059			0.054	0.054	0.054	0.054	0.054
Conditional Mean					0.028	0.022	0.017					0.001	-0.005	-0.006
Job FE	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	
Vacancy FE							✓							✓
Demographic controls		✓		✓		✓			✓		✓		✓	
Skill controls		✓		✓		✓			✓		✓		✓	

Notes: This table shows difference-in-difference estimates using the specification 1 but excluding the period 02/2020-05/2020 when take-up of anonymization was low due to randomization. Robust standard errors in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table F3: Main Results by Gender

	P(Interview)							P(Hired)						
	Foreign		Finnish		Pooled			Foreign		Finnish		Pooled		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Panel A: Women														
Treat X Post	0.046* (0.025)	0.047* (0.025)	-0.012 (0.018)	-0.012 (0.019)	-0.011 (0.018)	-0.011 (0.018)		0.033** (0.016)	0.034** (0.016)	-0.022** (0.011)	-0.022** (0.011)	-0.021** (0.010)	-0.021** (0.010)	
Foreign X Treat X Post					0.058** (0.025)	0.058** (0.025)	0.067*** (0.026)					0.052*** (0.018)	0.051*** (0.017)	0.053*** (0.020)
N	12,703	12,703	45,216	45,216	58,049	58,049	56,947	12,703	12,703	45,216	45,216	58,049	58,049	56,947
Foreign Mean	0.070	0.070			0.070	0.070	0.070	0.035	0.035			0.035	0.035	0.035
Finnish Mean			0.093	0.093	0.093	0.093	0.093			0.046	0.046	0.046	0.046	0.046
Job FE	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	
Vacancy FE							✓							✓
Demographic controls		✓		✓		✓			✓		✓		✓	
Skill controls		✓		✓		✓			✓		✓		✓	
Panel B: Men														
Treat X Post	0.073 (0.048)	0.068 (0.047)	0.019 (0.035)	0.014 (0.035)	0.022 (0.033)	0.020 (0.034)		0.045** (0.018)	0.039** (0.017)	-0.028 (0.024)	-0.031 (0.024)	-0.029 (0.023)	-0.029 (0.023)	
Foreign X Treat X Post					0.025 (0.046)	0.023 (0.046)	0.024 (0.050)					0.049* (0.029)	0.047 (0.029)	0.048* (0.029)
N	3,719	3,719	12,778	12,778	16,663	16,663	15,165	3,719	3,719	12,778	12,778	16,663	16,663	15,165
Foreign Mean	0.036	0.036			0.036	0.036	0.036	0.018	0.018			0.018	0.018	0.018
Finnish Mean			0.066	0.066	0.066	0.066	0.066			0.066	0.066	0.066	0.066	0.066
Job FE	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	
Vacancy FE							✓							✓
Demographic controls		✓		✓		✓			✓		✓		✓	
Skill controls		✓		✓		✓			✓		✓		✓	

Notes: This table shows difference-in-difference estimates using the specification 1 for foreign women in Panel A and foreign men in Panel B. Columns (1), (3), (5), (8), (10), and (12) control for job title fixed effects and provide the baseline estimates. Column (2), (4), (6), (9), (11), and (13) add controls for applicants age and gender. Columns (7) and (14) use vacancy fixed effects. The outcome means are pre-treatment period means for the treated group outcome variables. Robust standard errors in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table F4: Occupation Group by Close Control

Occupation group	Applicants			Vacancies		
	Distant	Close	Total	Distant	Close	Total
Vocational education teachers	1,962	0	1,962	164	0	164
Customer service and office tasks	1,689	5,145	6,834	26	104	130
Expert roles	209	9,938	10,147	7	323	330
Supervisory and managerial tasks	108	4,373	4,481	6	264	270
ICT tasks	14	1,335	1,349	1	93	94
Senior officials	525	0	525	21	0	21
Other tasks in education and teaching	591	6,750	7,341	46	468	514
Kitchen and cleaning sector tasks	3,493	0	3,493	376	0	376
Library, youth, sports, and cultural sector tasks	894	3,959	4,853	62	128	190
Transport, logistics, and real estate sector tasks	242	0	242	13	0	13
Practical nurses	3,795	0	3,795	825	0	825
Doctors and dentists	163	0	163	75	0	75
Primary and upper secondary school teachers	21,434	0	21,434	1,550	0	1,550
Nurses and health care professionals	2,970	0	2,970	602	0	602
Social workers	977	0	977	254	0	254
Other tasks in the social and healthcare sector	6,835	0	6,835	476	0	476
Technical and urban planning sector	2,632	0	2,632	213	0	213
Security, rescue, and preparedness tasks	298	0	298	25	0	25
Early childhood care assistants	5,780	0	5,780	883	0	883
Early childhood teachers	2,203	0	2,203	972	0	972
Early childhood teachers and social workers	44	0	44	18	0	18
Environmental and nature conservation sector tasks	280	0	280	14	0	14
Total	57,151	31,500	88,651	6,632	1,380	8,012

Table F5: Difference-in-Difference Estimates on Hiring and Calls to Interviews Close controls

	P(Interview)							P(Hired)						
	Foreign		Finnish		Pooled			Foreign		Finnish		Pooled		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Treat X Post	0.045** (0.022)	0.045** (0.023)	-0.008 (0.015)	-0.010 (0.016)	-0.007 (0.015)	-0.009 (0.015)		0.034** (0.014)	0.034** (0.014)	-0.024*** (0.008)	-0.025*** (0.009)	-0.023*** (0.008)	-0.024*** (0.008)	
Post X Close Control	-0.003 (0.014)	-0.004 (0.014)	-0.006 (0.012)	-0.007 (0.012)	-0.006 (0.011)	-0.007 (0.011)		0.016 (0.011)	0.015 (0.011)	-0.004 (0.004)	-0.004 (0.004)	-0.001 (0.004)	-0.002 (0.004)	
Foreign X Treat X Post					0.054** (0.021)	0.054** (0.021)	0.056*** (0.021)					0.053*** (0.016)	0.052*** (0.016)	0.049*** (0.017)
Foreign X Close X Post					0.004 (0.012)	0.005 (0.012)	0.014 (0.013)					0.008 (0.009)	0.009 (0.009)	0.011 (0.010)
N	16,593	16,593	58,102	58,102	74,802	74,802	73,877	16,593	16,593	58,102	58,102	74,802	74,802	73,877
Foreign Mean	0.063	0.063			0.063	0.063	0.063	0.031	0.031			0.031	0.031	0.031
Finnish Mean			0.086	0.086	0.086	0.086	0.086			0.051	0.051	0.051	0.051	0.051
Job FE	✓	✓		✓				✓	✓		✓			
Vacancy FE					✓	✓	✓			✓	✓	✓	✓	✓
Demographic controls		✓		✓		✓			✓		✓		✓	
Skill controls		✓		✓		✓			✓		✓		✓	

Notes: This table shows the difference-in-difference (Equation 1) and the triple difference estimates. Treat × Post is the estimated difference-in-differences coefficient, an intention-to-treat estimate of assignment to the anonymization pilot. In the subgroup columns (1–2 and 3–4) the sample is restricted to applicants from that group and Treat × Post gives the within-group treatment effect. In the pooled columns (5–7 and 12–14) all applicants are included and the triple interaction Foreign × Post × Treat (Panel A) tests whether the treatment effect differs between the two groups. Specifications with "Job FE" absorb job-title fixed effects; the last column of each outcome block uses vacancy fixed effects instead. Demographic controls are gender, age, and age squared. Skill controls are work experience, experience squared, and dummies for highest completed education level. Foreign Mean and Finnish Mean are pre-treatment means of the outcome among applicants from each subgroup in treated vacancies; in the pooled columns both are reported. Robust standard errors in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table F6: Differences-in-Differences and Triple Difference Estimates Using Alternative Measures of Ethnic Identity

	P(Interview)							P(Hired)						
	Foreign		Finnish		Triple difference			Foreign		Finnish		Triple difference		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Panel A: Language (Nonnative)														
Treat X Post	0.039** (0.019)	0.040** (0.019)	0.004 (0.015)	0.001 (0.015)	0.005 (0.015)	0.004 (0.015)		0.036*** (0.009)	0.037*** (0.010)	-0.015** (0.006)	-0.017** (0.007)	-0.015** (0.006)	-0.015** (0.007)	
Foreign X Treat X Post					0.019 (0.021)	0.022 (0.021)	-0.001 (0.020)					0.041*** (0.012)	0.041*** (0.012)	0.022* (0.012)
N	8,242	8,242	66,407	66,407	74,802	74,802	73,877	8,242	8,242	66,407	66,407	74,802	74,802	73,877
Foreign Mean	0.034	0.034			0.034	0.034	0.034	0.008	0.008			0.008	0.008	0.008
Finnish Mean			0.086	0.086	0.086	0.086	0.086			0.051	0.051	0.051	0.051	0.051
Conditional Mean					-0.043	-0.033	-0.045					-0.037	-0.031	-0.035
Panel B: Full Name														
Treat X Post	0.034 (0.030)	0.035 (0.030)	0.006 (0.014)	0.004 (0.014)	0.006 (0.014)	0.005 (0.014)		0.041*** (0.012)	0.042*** (0.013)	-0.013** (0.006)	-0.014** (0.006)	-0.012** (0.006)	-0.013** (0.006)	
Foreign X Treat X Post					0.028 (0.029)	0.031 (0.029)	0.009 (0.029)					0.044*** (0.013)	0.044*** (0.013)	0.025* (0.014)
N	4,786	4,786	69,855	69,855	74,802	74,802	73,877	4,786	4,786	69,855	69,855	74,802	74,802	73,877
Foreign Mean	0.039	0.039			0.039	0.039	0.039	0.000	0.000			0.000	0.000	0.000
Finnish Mean			0.083	0.083	0.083	0.083	0.083			0.050	0.050	0.050	0.050	0.050
Conditional Mean					-0.049	-0.036	-0.049					-0.053	-0.038	-0.051
Panel C: First Name														
Treat X Post	0.048 (0.031)	0.050 (0.031)	0.002 (0.014)	0.000 (0.014)	0.003 (0.014)	0.001 (0.014)		0.056*** (0.014)	0.058*** (0.014)	-0.018*** (0.006)	-0.019*** (0.007)	-0.017*** (0.006)	-0.018*** (0.007)	
Foreign X Treat X Post					0.048* (0.029)	0.049* (0.029)	0.026 (0.028)					0.072*** (0.015)	0.070*** (0.015)	0.062*** (0.015)
N	7,403	7,403	67,241	67,241	74,802	74,802	73,877	7,403	7,403	67,241	67,241	74,802	74,802	73,877
Foreign Mean	0.045	0.045			0.045	0.045	0.045	0.009	0.009			0.009	0.009	0.009
Finnish Mean			0.084	0.084	0.084	0.084	0.084			0.051	0.051	0.051	0.051	0.051
Conditional Mean					-0.046	-0.042	-0.042					-0.048	-0.043	-0.044
Panel D: Last Name														
Treat X Post	0.042* (0.025)	0.044* (0.025)	0.001 (0.015)	-0.001 (0.015)	0.002 (0.015)	0.001 (0.015)		0.019 (0.015)	0.020 (0.015)	-0.015** (0.007)	-0.016** (0.008)	-0.014* (0.007)	-0.015** (0.007)	
Foreign X Treat X Post					0.031 (0.022)	0.034 (0.022)	0.031 (0.022)					0.026* (0.016)	0.026* (0.015)	0.017 (0.017)
N	10,962	10,962	63,690	63,690	74,802	74,802	73,877	10,962	10,962	63,690	63,690	74,802	74,802	73,877
Foreign Mean	0.051	0.051			0.051	0.051	0.051	0.023	0.023			0.023	0.023	0.023
Finnish Mean			0.086	0.086	0.086	0.086	0.086			0.051	0.051	0.051	0.051	0.051
Conditional Mean					-0.029	-0.026	-0.037					-0.025	-0.022	-0.029
Job FE	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	
Vacancy FE							✓							✓
Demographic controls		✓		✓		✓			✓		✓		✓	
Skill controls		✓		✓		✓			✓		✓		✓	

Notes: This table shows difference-in-difference estimates using the specification 1 using alternative measures of ethnic identity. Panel A reports the results for nonnative language, Panel B for full name, Panel C for first name, and Panel D for last name. Columns (1), (3), (5), (8), (10), and (12) control for job title fixed effects and provide the baseline estimates. Column (2), (4), (6), (9), (11), and (13) add controls for applicants age and gender. Columns (7) and (14) use vacancy fixed effects. The outcome means are pre-treatment period means for the treated group outcome variables. Robust standard errors in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

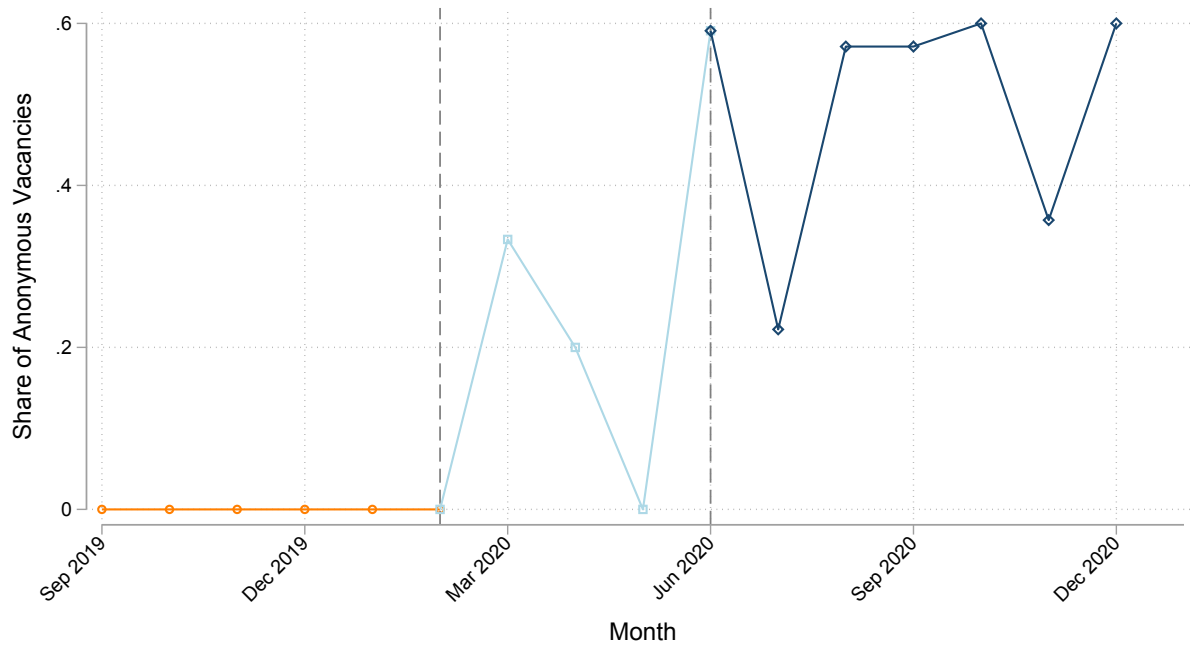
Table F7: Sensitivity of Hiring Outcomes to Sample Definition: Dropping Occupation Groups and Administrative Units

Panel A: Dropping one occupation group at a time				
Excluded occupation group	Take-up	Interview	Hired	N
	(1)	(2)	(3)	(4)
Vocational education teachers	0.444	0.0506** (0.0210)	0.0493*** (0.0156)	72,896
Customer service and office tasks	0.494	0.0741* (0.0379)	0.0839*** (0.0250)	69,897
Expert roles	0.426	0.0475** (0.0205)	0.0477*** (0.0160)	67,789
Supervisory and managerial tasks	0.442	0.0536** (0.0212)	0.0493*** (0.0158)	71,255
ICT tasks	0.439	0.0518** (0.0210)	0.0494*** (0.0156)	73,866
Senior officials	0.444	0.0521** (0.0210)	0.0498*** (0.0156)	74,462
Other tasks in education and teaching	0.239	0.0664*** (0.0204)	0.0566*** (0.0161)	68,251
Kitchen and cleaning sector tasks	0.444	0.0466** (0.0211)	0.0443*** (0.0156)	71,642
Library, youth, sports, and cultural sector tasks	0.650	0.0274 (0.0238)	0.0267 (0.0164)	72,155
Transport, logistics, and real estate sector tasks	0.444	0.0521** (0.0210)	0.0498*** (0.0156)	74,589
Practical nurses	0.444	0.0523** (0.0211)	0.0489*** (0.0156)	71,522
Doctors and dentists	0.444	0.0519** (0.0210)	0.0496*** (0.0156)	74,656
Primary and upper secondary school teachers	0.444	0.0548** (0.0214)	0.0499*** (0.0160)	53,616
Apprenticeships	0.444	0.0520** (0.0210)	0.0496*** (0.0156)	74,789
Nurses and health care professionals	0.444	0.0524** (0.0211)	0.0503*** (0.0156)	72,081
Social workers	0.444	0.0511** (0.0210)	0.0486*** (0.0156)	73,860
Other tasks in the social and healthcare sector	0.444	0.0526** (0.0211)	0.0497*** (0.0157)	69,166
Technical and urban planning sector	0.444	0.0518** (0.0210)	0.0496*** (0.0156)	72,823
Security, rescue, and preparedness tasks	0.444	0.0518** (0.0210)	0.0495*** (0.0156)	74,631
Early childhood care assistants	0.444	0.0511** (0.0211)	0.0491*** (0.0157)	69,298
Early childhood teachers	0.444	0.0547*** (0.0210)	0.0523*** (0.0156)	72,717
Early childhood teachers and social workers	0.444	0.0521** (0.0210)	0.0498*** (0.0156)	74,752
Environmental and nature conservation sector tasks	0.444	0.0520** (0.0210)	0.0498*** (0.0156)	74,645
Panel B: Dropping one administrative unit at a time				
Excluded administrative unit	Take-up	Interview	Hired	N
	(1)	(2)	(3)	(4)
Helsinki City Transport	0.439	0.0535** (0.0214)	0.0499*** (0.0159)	73,331
The Helsinki City Executive Office	0.438	0.0552*** (0.0211)	0.0514*** (0.0157)	70,714
Education Division	0.215	0.0728*** (0.0214)	0.0637*** (0.0170)	33,568
Culture and Leisure Division	0.661	0.0282 (0.0238)	0.0259 (0.0164)	71,962
Urban Environment Division	0.444	0.0519** (0.0210)	0.0501*** (0.0156)	71,956
Service Centre	0.424	0.0425** (0.0205)	0.0410** (0.0161)	71,192
Rescue Department	0.444	0.0515** (0.0210)	0.0492*** (0.0156)	74,559
Social Services and Health Care Division	0.500	0.0681* (0.0368)	0.0804*** (0.0242)	57,006
Construction Services Public Utility	0.444	0.0520** (0.0210)	0.0497*** (0.0156)	74,680
Financial Management Services	0.444	0.0522** (0.0210)	0.0500*** (0.0156)	74,521
Audit Department	0.444	0.0520** (0.0210)	0.0497*** (0.0156)	74,802
Occupational Health Centre	0.444	0.0520** (0.0210)	0.0497*** (0.0156)	74,531

Notes: Each row re-estimates the triple-difference specification of Table 2, Columns (5) and (12), after removing the indicated occupation group (Panel A) or administrative unit (Panel B). Columns (2) and (3) report the Foreign x Treat x Post coefficient on the probability of being called to an interview and of being hired, with robust standard errors beside each estimate. All specifications include job-title and month fixed effects. Take-up is the share of treated hiring processes that were anonymized in the remaining sample. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

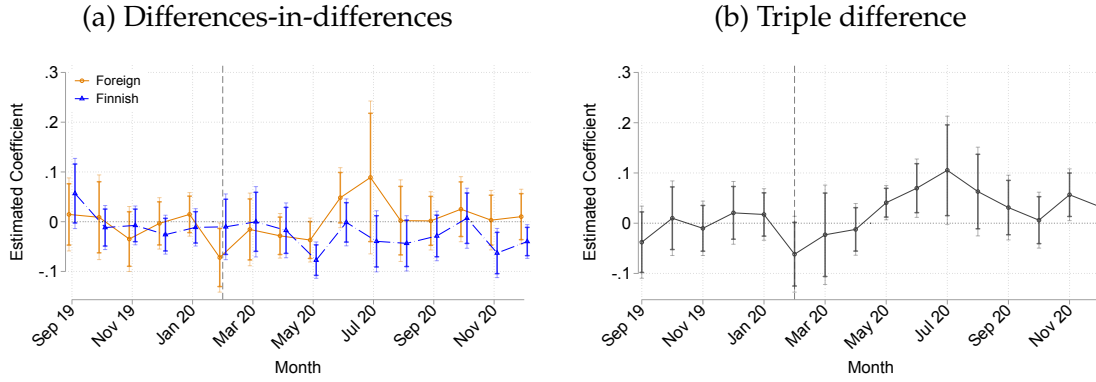
G Figures

Figure G1: Anonymization in Treated Job Titles

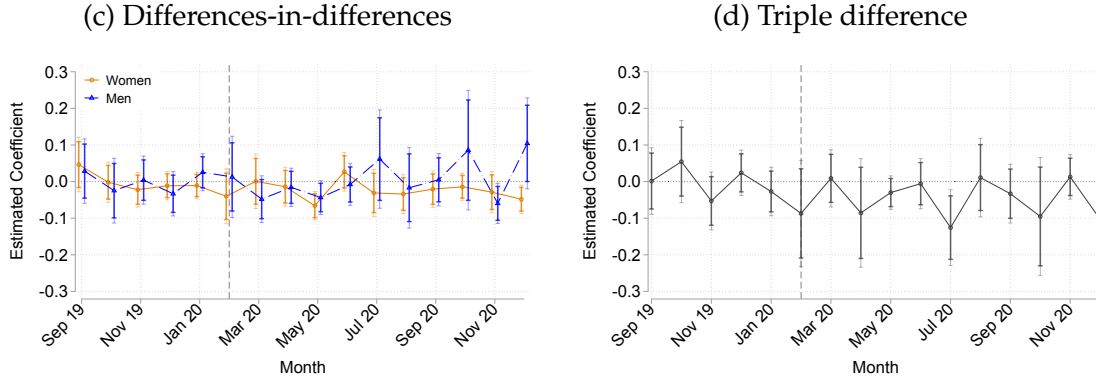


Notes: This figure shows the share of anonymized recruitment processes at the City of Helsinki around the pilot period in the treatment group. The dashed vertical line marks the beginning of the pilot program. The dashed horizontal line shows the average during the whole pilot period.

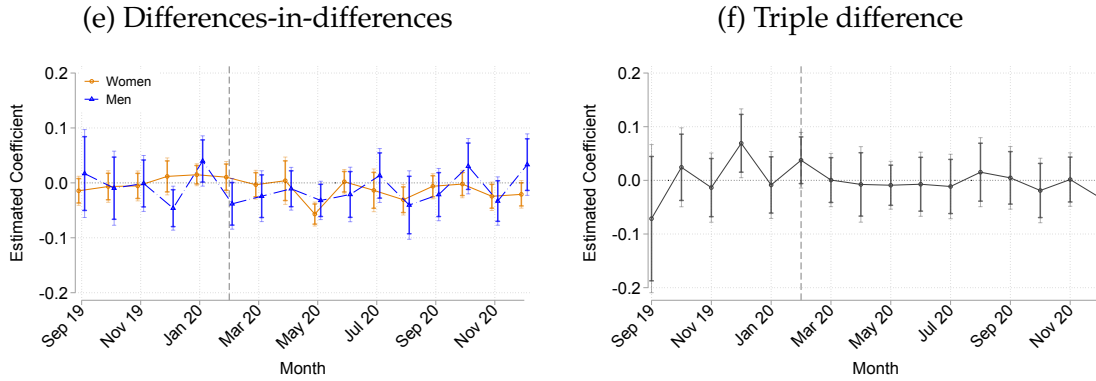
Figure G2: Event Study Estimates
Calls to Interviews by Immigrant Background



Calls to Interviews by Gender



Hiring by Gender



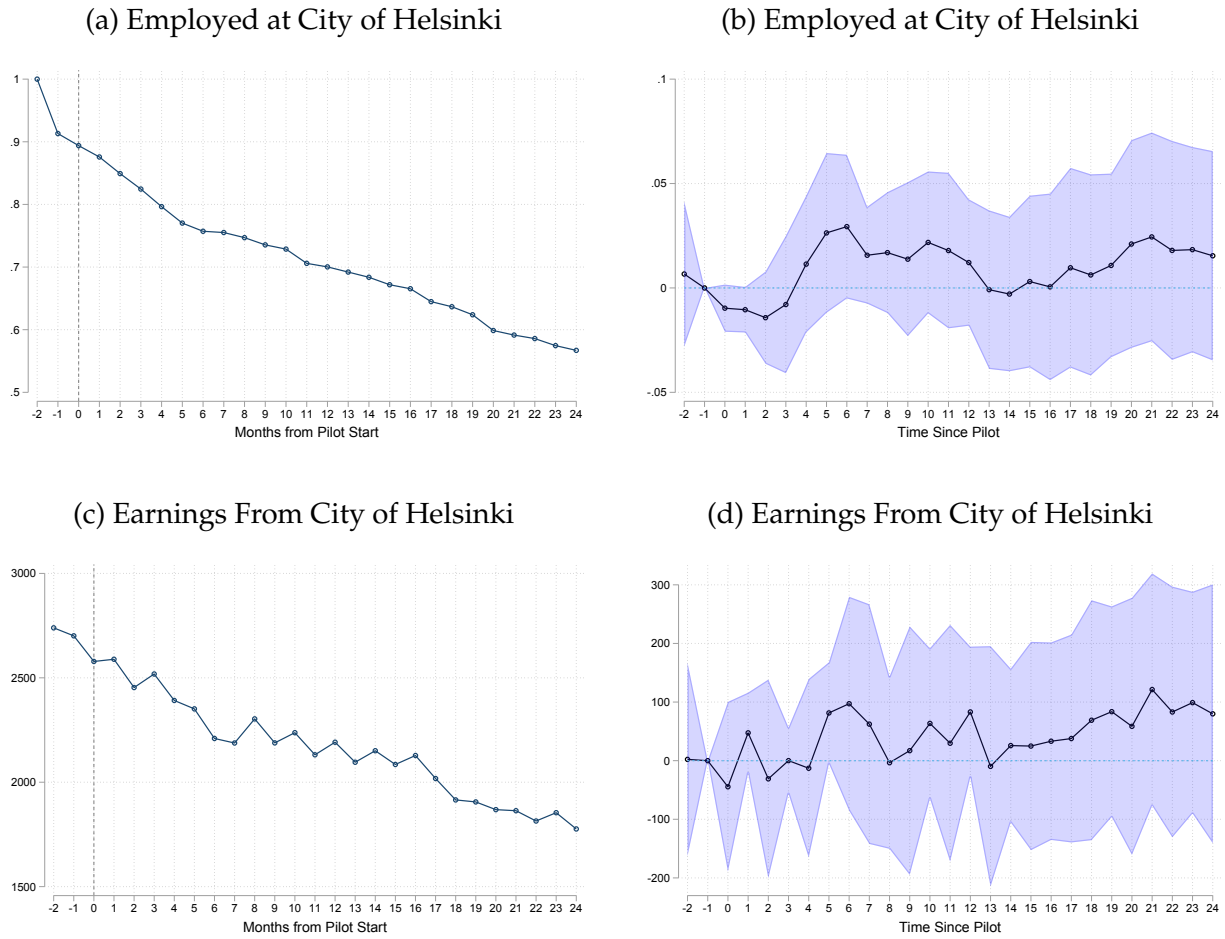
Notes: These figures show the event-study estimates on calls to interviews and hiring decisions. Panel (a) shows the estimates separately for applicants with foreign (yellow) and Finnish (blue) names from the following specification: $Y_{ijt} = \alpha + \sum_{k=-5}^{11} \beta_k \text{Treated}_j \times \mathbf{1}\{t - \text{Feb}2020 = k\} + \lambda_t + \theta_j + \varepsilon_{ijt}$, where we restrict $\sum_{k=-5}^{-1} \beta_k = 0$. Panel (b) shows analogous estimates for the gap between applicants with foreign relative to applicants with Finnish names from a triple-diff specification. Panels (c)–(f) show similar estimates on calls to interviews and hiring decisions by gender. All panels show 90 and 95 percent confidence intervals around point estimates.

Figure G3: Employment and Wages Around Job Start



Notes: These figures show work status and earnings around the job start month as defined by the first month the worker received a positive wage. Panels A and C show the average employment at and earnings from the City of Helsinki for workers who started their job in a given month. Panels B and D show the difference-in-difference estimates comparing employment and earnings trajectories of workers hired in the treatment group vs the control group processes. Each point estimate comes from a separate regression. Shaded areas show 95% confidence intervals.

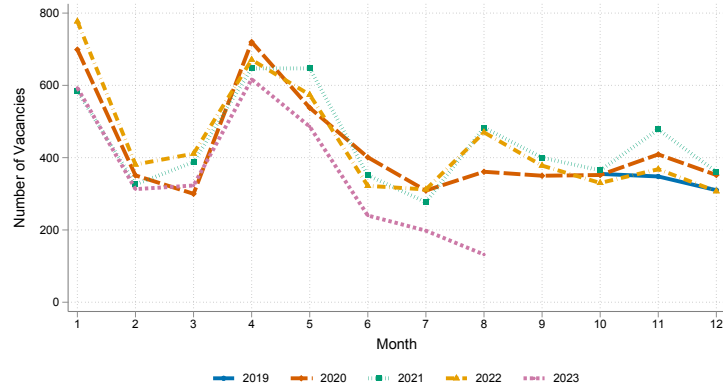
Figure G4: Incumbent Workers' Employment and Earnings around the Pilot



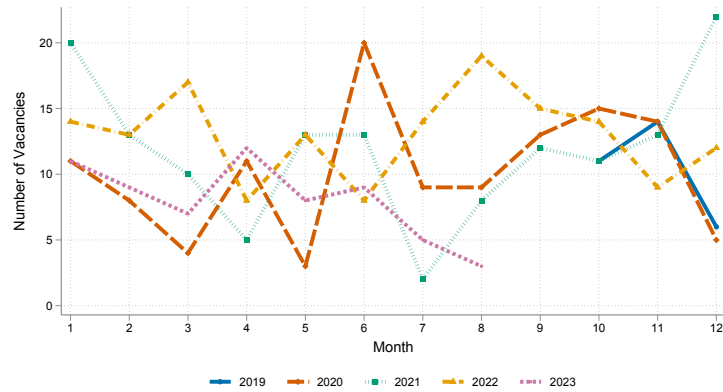
Notes: This figure shows labor market outcomes for workers who were already employed at the City of Helsinki in December 2019, before the pilot started. Panel A shows the probability that these workers continued employment at the City of Helsinki. Panel B shows event-study estimates on the differences in employment trajectories of workers in the treated vs control job titles. Panel C and D show similar figures but for monthly earnings. Shaded areas show 95% confidence intervals.

Figure G5: Monthly Vacancies

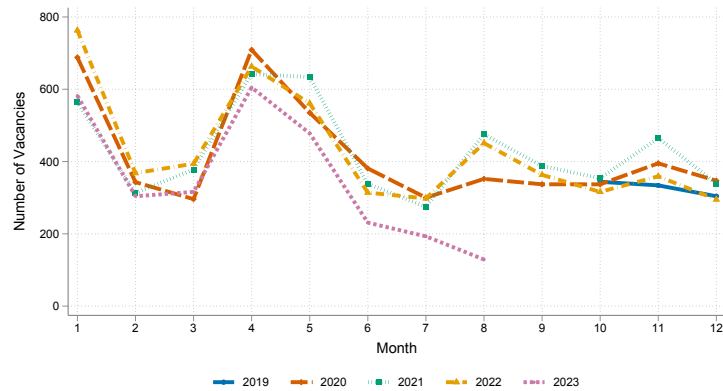
(a) All vacancies



(b) Treated vacancies

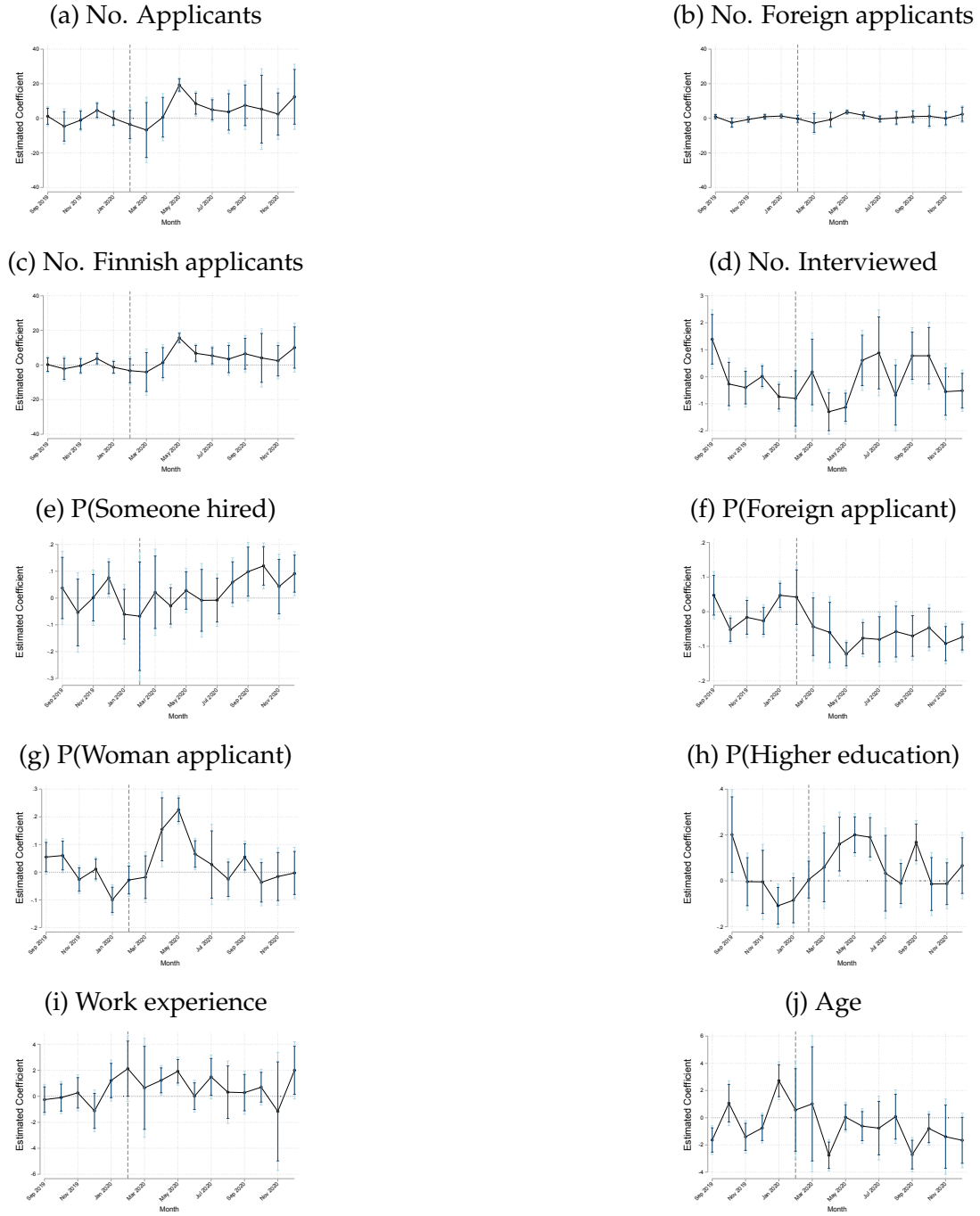


(c) Control vacancies



Notes: The figure shows the number of monthly vacancies in the years 2019-2023. Panel (a) shows the total number of vacancies. Panel (b) describes the treated job titles and Panel (c) shows vacancies in the control job titles.

Figure G6: Event-Study Estimates of Applicant Characteristics



Notes: These figures show the event-study estimates for the outcomes presented in Table 3. Panels (a)–(e) show the application process outcomes. Panels (f)–(j) show the applicant outcomes. The figures show the estimates from the following specification: $Y_{ijt} = \alpha + \sum_{k=-5}^{11} \beta_k \text{Treated}_j \times \mathbf{1}\{t - \text{Feb2020} = k\} + \lambda_t + \theta_j + \varepsilon_{ijt}$, where we restrict $\sum_{k=-5}^{-1} \beta_k = 0$. Figures show 90 and 95 percent confidence intervals around point estimates.