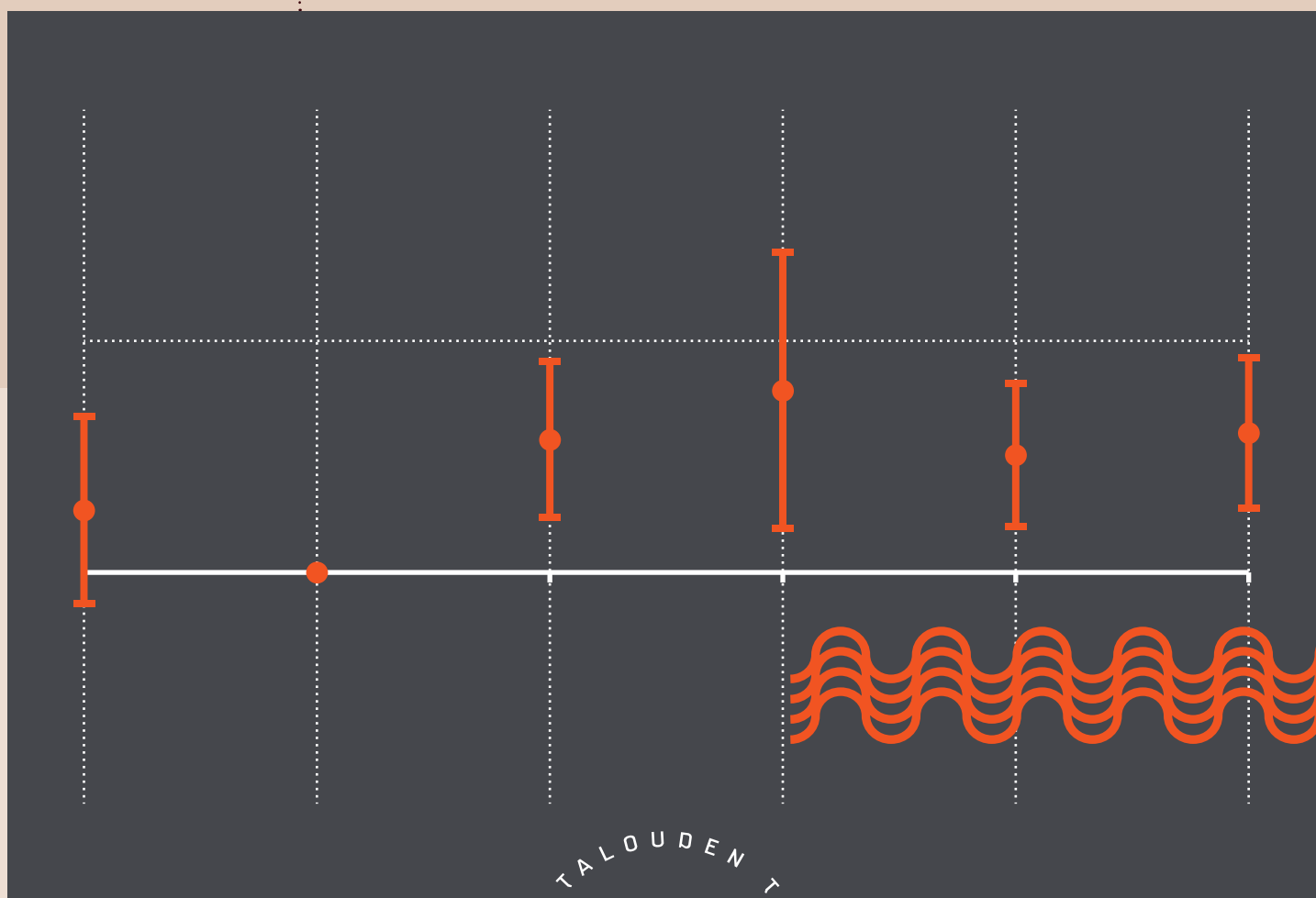


Ohto Kanninen**
Sanni Kiviholma***
Tuomo Virkola****

TYÖPAPEREITA / WORKING PAPERS

Anatomy of An Anonymous Hiring Pilot*

338



TYÖN JA TALOUDEN TUTKIMUS
EST 1971
Labore

* We thank participants in Labore, VATT and FDPE seminars and Hanna Pesola for helpful comments. We also thank the City of Helsinki for providing us ample information on the details of the pilot and access to the administrative data. We are grateful for Ami-säätiö for funding. Any errors are ours.

** Labour Institute for Economic Research LABORE
(Helsinki, Finland), ohto.kanninen@labore.fi

*** University of Jyväskylä & Labour Institute for
Economic Research LABORE (Helsinki, Finland),
sanni.kiviholma@labore.fi

**** VATT Institute for Economic Research,
tuomo.virkola@vatt.fi.

ABSTRACT

Anonymous hiring restricts the information set available to hiring managers and can create a trade-off between discrimination and hiring efficiency. To study this potential trade-off, we leverage a pilot by the City of Helsinki during which hiring to some but not all job titles were subject to anonymization. We find a positive effect on hiring applicants with foreign names but no effect on job duration or wages. Announcing anonymous hiring increases job applications but this is not driven by applicants with foreign-sounding names. Instead, we observe more women applying. We find evidence that once hiring managers are allowed to opt in to anonymous hiring (vs mandated), this induces self-selection which could have important

consequences for the effectiveness of the policy and explain why our results differ from some previous research. Results suggest that anonymization can work as an effective tool to fight against discrimination at least when it is not an opt-in policy.

JEL Classification: J71, J78

Keywords:

anonymous applications, discrimination, employment

TIIVISTELMÄ

Tutkimme Helsingin kaupungin anonyymin rekrytoinnin pilotin avulla anonymisoinnin vaikutusta syrjintään ja palkattujen menestymiseen töissä. Havaitsemme, että henkilöitä, joilla on ulkomaalaisen kuuloinen nimi, palkattiin enemmän. Anonymisointi ei sen sijaan vaikuta palkattujen sukupuoleen. Anonymisoinnin potentiaalinen haitta on se, että tiedon rajaaminen heikentää mahdollisuuksia palkata sopivin hakija. Emme kuitenkaan havaitse vaikutusta menestymiseen töissä työskentelyjakson kestolla tai palkoilla mitattuna. Rekrytointiprosessin anonymisointi lisäsi hakijoiden määrä kokonaisuudessaan, erityisesti naishakijoiden mTutkimme Helsingin kaupungin anonyymin rekrytoinnin pilotin avulla anonymisoinnin vaikutusta syrjintään ja palkattujen menestymiseen töissä. Havaitsemme, että henkilöitä, joilla on ulkomaalaisen kuuloinen nimi, palkattiin enemmän. Anonymisointi ei sen sijaan vaikuta palkattujen

sukupuoleen. Anonymisoinnin potentiaalinen haitta on se, että tiedon rajaaminen heikentää mahdollisuuksia palkata sopivin hakija. Emme kuitenkaan havaitse vaikutusta menestymiseen töissä työskentelyjakson kestolla tai palkoilla mitattuna. Rekrytointiprosessin anonymisointi lisäsi hakijoiden määrä kokonaisuudessaan, erityisesti naishakijoiden määrä kasvoi. Pilotin jälkeen, kun esihenkilöt saivat itse valita anonymisoivatko he rekrytoinnin, havaitsemme, että erilaiset henkilöt päätyvät anonymisoimaan, kuin kokeilun aikana, jolloin se oli pakollista. Tulokset antavat viitteitä, että anonyymi rekrytointi voi olla tehokas työkalu syrjinnän vähentämiseksi, vaikka on kiinnitettävä huomiota, kuinka sen käyttöönotto toteutetaan.

Avainsanat:

anonyymit hakemukset, syrjintä, työllisyys

1 Introduction

Social science literature demonstrates that discrimination based on ethnic background exists in the labor markets.¹ Correspondence studies have provided evidence that a job seeker’s name and the group membership it signals can affect callback rates to job applications.² This evidence has motivated proposals to introduce anonymous hiring policies in which information about applicants’ names and group membership is removed. This should make it harder for employers to directly discriminate in the early stages of the hiring process. However, limiting information about a job candidate can affect hiring and the quality of the hiring process if the name or group membership carries over some information about the workers’ expected productivity or suitability for the job. Yet, there is not much evidence on these potential costs and benefits of anonymous hiring.

To make progress in understanding these potential trade-offs, we leverage a City of Helsinki recruitment pilot program. During the pilot, new hires to a pre-defined set of job titles were subjected to an anonymous hiring process in which hiring managers were not able to see identifying information about a job candidate before calls to interviews were sent out. This allows us to use a difference-in-difference type framework comparing hiring across job titles over time with a plausibly exogenous change in anonymization which is not driven by selection at the hiring manager level. For the empirical analysis, we use data on the universe of job applicants, calls to interviews, and new hires at the City of Helsinki. 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. Further, 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 vs. those who were not. This allows us to test how the hiring process affects the quality of matches between jobs and workers. Finally, leveraging a switch in the anonymization policy after the pilot and data on hiring manager characteristics, we study how selection into anonymizing depends on whether it is mandated or voluntary.

In our baseline analysis, we find evidence that the City of Helsinki recruitment pilot increased the likelihood that applicants with foreign-sounding names were hired. Our

¹See, f.e. [Lang and Lehmann \(2012\)](#) and [Neumark \(2018\)](#).

²See [Bertrand and Mullainathan \(2004\)](#), [Pager \(2003\)](#) for seminal studies using an experimental design and [Quillian et al. \(2017\)](#) for a meta-analysis and trends on discrimination in the U.S. and [Neumark \(2018\)](#) for a more comprehensive review of the literature. See [Ahmad \(2020\)](#) for similar evidence in the Finnish context for job applicants with foreign-sounding names. See [Kessler et al. \(2019\)](#) for a measure of discrimination that is not based on deception.

estimates suggest that the policy eliminated the hiring gap between applicants with foreign and Finnish names. The effect does not seem to be driven only by a change in call to interviews. Instead, we see that the effect on hiring decisions is not smaller than the effect on calls to interviews. Thus, it seems that the policy was effective in promoting the employment prospects of an applicant group that was primarily targeted by the policy. On the other hand, we do not observe that anonymization would have affected the gender composition of calls to interviews or hiring decisions.

We also find new evidence that anonymizing job processes and advertising it 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 primarily driven by applications from individuals with foreign-sounding names who were targeted with the policy and who ex-ante could have benefitted the most. Instead, we see that the share of applicants with foreign-sounding names decreases and the share of women increases. 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 both among individuals with foreign-sounding names as well as for women, although through different mechanisms. However, in our case, these results on the total effects remain imprecise and not statistically significant at standard levels.

Finally, we find that when anonymization is made voluntary for hiring managers, it induces potentially important selection that could affect the size or even the sign of the effects of the policy. During the pilot, when the anonymization is mandated in pre-determined job titles, individual manager characteristics do not predict the use of anonymization. After the pilot, however, manager characteristics become correlated with anonymization decisions. This type of selection could explain why our baseline results differ from those estimated in the previous literature (see [Behaghel et al. \(2015\)](#)). However, our data does not allow us to directly say much about the potential direction of this selection.

Our study relates to a large literature on discrimination and empirical research establishing that there exists discrimination based on the names of job applicants and their group membership. The empirical research compellingly demonstrates that employers use names as informative signals about job applicants in the early stages of recruitment.³

³See, f.e. [Lang and Lehmann \(2012\)](#) and [Neumark \(2018\)](#).

Our results are consistent with this view. Less is known about how the removal of identifying information affects hiring decisions and the quality of the hiring process.⁴ 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 has provided some evidence in this direction.⁵

Our study is closely linked to literature on anonymous hiring policies which have so far provided somewhat conflicting results.⁶ These papers are mostly concerned with the early recruitment stages or find evidence of reduced discrimination in calls to interviews but not on final hiring outcomes. For example, [Åslund and Skans \(2012\)](#) find in a context which is similar to ours that anonymization leveled the playing field in calls to interviews, but the effects on hiring decisions only carry over for women but not for immigrant applicants. We complement this previous research by analyzing the whole recruitment process, starting from applications to interviews and hiring decisions, and also analyzing the effects on job duration and wages on the job. This allows us to better assess potential trade-offs that are induced by the policy and which employers need to consider when making decisions about their recruitment processes.

The effects of anonymization might also depend importantly on the way it is implemented and on the types of firms or institutions that take up the policy. A prominent paper using a credible research design finds evidence that in some contexts anonymization can actually be counterproductive to minorities that are the target group of the policy ([Behaghel et al. \(2015\)](#)). The authors suggest that this can be driven by the fact that firms participating in the study were those who, in essence, were positively discriminating against minority applicants. They hypothesize that the effects could have been different if anonymization had been mandated. This is what our setting allows us to test and thus our setting provides evidence for policy makers considering to mandate anonymization.

Finally, our study also relates to some recent work on trying to understand how different recruitment techniques affect job seekers' application decisions. When a potential applicant observes that recruiting is anonymized she may or may not be encouraged to

⁴See [Agan and Starr \(2018\)](#), [Hoffman et al. \(2018\)](#), [Kolev et al. \(2019\)](#)

⁵See [Agan and Starr \(2018\)](#), [Doleac and Hansen \(2018\)](#) and [Veuger and Shoag \(2019\)](#) on evidence in the case of suppressing information on criminal records and how it affects employment outcomes of the exposed groups.

⁶See [Goldin and Rouse \(2000\)](#), [Behaghel et al. \(2015\)](#), [BETA \(2017\)](#), [Krause et al. \(2012b\)](#), [Krause et al. \(2012a\)](#), [Åslund and Skans \(2012\)](#), [Public Service Commission of Canada \(2018\)](#) for evaluations of the effects of anonymization in the recruiting context.

apply because a) she might perceive her chances of acceptance to change or b) she might expect that if hired or not hired, it could be because of her background characteristics, e.g. due to an affirmative action policy. For example, [Leibbrandt and List \(2018\)](#) find evidence that racial minorities are less likely to apply to jobs that are advertized with equality-of-opportunity statements. They interpret this finding to be due to “tokenism” where minorities perceive their recruitment to be due to other factors than their merits. We provide new evidence on how a specific anti-discrimination policy might affect application decisions.

This paper is constructed as follows. Section 2 describes the anonymous hiring pilot. Section 3 describes the data. Section 4 describes our research design. Sections 5 and 6 present our results.

2 The City of Helsinki Anonymous Hiring Pilot

In our analysis, we study an anonymous hiring pilot program at the City of Helsinki. In December 2018, the City Council enacted a motion to introduce the possibility of conducting hiring processes anonymously. The aim was to "advance the opportunities of individuals that face possible inadvertent discrimination".⁷ This was motivated in part by i) the observation that some population groups (especially individuals with immigrant or foreign background) were underrepresented among the city’s employees and ii) by the existing evidence that some applicant groups are likely to face a penalty in the early recruiting stages.

During the pilot, all identifiable information about the candidate was hidden from the application files before invitations to job interviews were sent out. In practice, the HR department forwarded summary files of applicants to the hiring managers. These files included information on previous job experience, education, and degree titles as well as motivation letters. In contrast, no information on the applicant’s name, address, age, language skills, or educational institution was included. The length of previous work experience was shown but the exact dates were not. Hiring managers would then select candidates based on these summary files. After sending out calls to interviews, the full applicant files became available to hiring managers and all interviews and hiring decisions are done in a standard non-anonymous way.

The anonymous hiring pilot started in February 2020 and continued until December

⁷The motion (in Finnish): <https://dev.hel.fi/paatokset/asia/hel-2018-006254/kvsto-2018-22/>

2020. For the duration of the pilot, the City of Helsinki selected 13 job titles that would go through an anonymous hiring process. This creates the variation that we leverage in our empirical setting. These job titles include f.e. administrative assistants and secretaries, youth workers, coordinators, and planners. Importantly, the selected job titles do not include some of the most common job titles where the city faced the most difficulties in finding suitable candidates (such as nurses or practical nurses). This affects the interpretation of the pilot as it is mostly informative about hiring in a context where there are ex-ante no significant hiring difficulties or labor scarcity. Further, the initial plan was to randomize anonymization among job openings that belonged to the selected job titles. However, due to fewer than expected job openings to those job titles, and due to the observation that a surprisingly low share of the randomization allocated job openings to the control group, the city decided to stop randomizing and subjected all job openings in the selected job titles to an anonymous hiring process.

3 Data

We use administrative data from the City of Helsinki recruitment system that covers the universe of job openings and job applications for the period between September 2019 and May 2021. We observe all basic information submitted by applicants to the recruitment system, including job experience, education, and language skills. Our data also includes all the information contained in the job ads as observed by the job applicants, including the job title and whether or not the recruitment process is anonymous or not. 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 allows us to study recruiting choices with essentially the same information set that was available to hiring managers before sending invitations for job interviews.

For our primary analysis, we construct name-based measures of perceived ethnicity and gender that are not directly observed in our data. For this purpose, we use 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 on this list into either Nordic or non-Nordic sounding names.⁸ We instructed the research assistant to classify names based on first impressions to simulate the environment in which a hiring

⁸A majority of Finnish names are either Finnish or Swedish sounding and separating Nordic from the two is in most cases not feasible or meaningful.

manager evaluates job candidates. We then match this subjective classification to job applicants in our primary data. Second, to construct a measure of applicants' gender, we calculate the share of Finns who are women and have a specific name. If there are both men and women with the same name, we classify a name based on the majority gender with that name. Constructing a name-based gender measure is useful as it reflects also what the hiring managers observe and what they can infer, but also because in many City of Helsinki job openings gender has not been collected directly from applicants.

To measure wages and job duration, we match all new hires to administrative payroll data using names and birth dates (see the Appendix for details). The payroll data contains information on the monthly earnings at the individual level. These data correspond to standard pay-slip information. We classify a worker as employed at the City of Helsinki whenever the monthly earnings are positive. 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 at the time of hiring, we define the job start using the first earnings from a new job spell observed in the payroll data.

In the empirical analysis, we concentrate on job openings with one vacancy. This covers more than 85 percent of all job openings and allows us to exclude job openings that could include multiple job titles and could create some measurement error in whether or not those hiring processes should be in the treatment or control groups. In the individual-level analysis, we also use data on applicants whose complete background information (date of birth, education, and job experience) we observe. This makes sure that we are using the same sample in all the analyses with or without individual-level controls. We also need birth dates to later match applicants to payroll data so this creates consistency in these separate analyses. However, the estimates on the main hiring outcomes do not qualitatively depend on whether we restrict the sample to these individuals.

Table 1 shows the number of applicants, vacancies, and hires in the pre-pilot and pilot periods in the treated and control job titles. In total, our sample covers more than eight thousand job openings between September 2019 and December 2020 at the City of Helsinki. This resulted in more than 90 thousand applications and three and a half thousand recruitments. However, only a relatively small share of all the job openings fall into the category of the treated job titles. During the pilot, there were 99 job openings in the treatment group which attracted 2773 applicants and resulted in 62 hires.

Table 2 describes the pre-treatment characteristics of job openings, applicants, and recruits in our sample. On average, there are about 30 applicants per job opening. A

majority of the applicants are women which is consistent with the observation that many occupations in the public sector have been traditionally predominantly female. Around 7 percent of the applicants have a foreign-sounding first name as per our measure explained above. This roughly matches the share of applicants who do not report Finnish or Swedish as their first language. The average applicants are about 38 years old and a majority have a bachelor’s degree and about ten years of previous work experience. On average the applicants have similar observable characteristics but have somewhat more previous work experience than applicants who become recruited.

4 Empirical strategy

Our research design uses the variation in the anonymized recruiting process induced by the pilot program. Following the literature, our aim is to estimate how the chances of becoming interviewed or hired differ for applicants in an anonymized vs. standard process. Thus, in our baseline analysis, we leverage information at the applicant level and the specification takes the form

$$Y_{ijt} = \alpha + \beta D_i^g * \text{Treated}_j * \text{Post}_t + \delta X_{ijt} + \lambda_t + \theta_j + \varepsilon_{ij} \quad (1)$$

where Y_{ijt} is the outcome of applicant i who applied to a job of job title j at time t , D_i^g is a dummy for applicant’s group membership $g \in \{\text{Foreign-sounding name, woman}\}$, Treated_j is a dummy for the job opening being among the treated job titles, Post_t is a dummy for the pilot duration and X_{ijt} is a set of controls including pairwise interaction terms of D_i^g , Treated_j and Post_t as well as potential job and applicant level characteristics. To the extent that recruitment outcomes in the treated and untreated programs would have evolved in parallel in the absence of the pilot program, this specification yields an unbiased estimate of the intention to treat $\hat{\beta}$.

To get a more comprehensive understanding of how anonymization affects hiring outcomes, we need to consider both the likelihood that an individual from a certain group is hired as well as the share of applicants from that group that decide to apply for that job. Thus, we complement our baseline analysis by leveraging variation purely between job openings in a standard difference-in-difference framework. We estimate models of the form

$$Y_{jt} = \alpha + \beta * \text{Treated}_j * \text{Post}_t + \lambda_t + \theta_j + \varepsilon_{ij} \quad (2)$$

where Y_{jt} is the average outcome for individuals who applied or were hired to job opening j . Now, depending on the outcome variable, $\hat{\beta}$ is informative about the effect on applications as well as on the total effect of hiring individuals from a specific group. We also use an analogous framework to study the effects on wages and employment for workers hired to job opening j . In both of these cases, the identifying assumption remains the same in the absence of the treatment, hiring in the treatment and control job titles would have evolved in parallel 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 2 shows the average characteristics of applicants and recruits in the control and treated job titles. It shows that both the applicants and recruits look on average somewhat different between the two groups. For example, job openings in the treatment group attract more applicants, more of whom have foreign-sounding names and who are less likely to have a bachelor's degree. Otherwise, the differences are small in magnitude. Yet, these observed differences highlight the importance of accounting for differences in the hiring outcomes prior to the pilot. This is what our research design based on the difference-in-differences does.

5 Main Results

In this section, we present our main results. We start by describing the timing and implementation of the hiring pilot. We then present our main results where we compare hiring outcomes among applicants in the treatment group vs. the control group. Next, we analyze the effects on application choices as well as the total effects on hiring which takes both into account. Finally, we discuss results on employment and wages after the job starts to provide evidence of the worker-job match quality.

5.1 Anonymization Among Treated Job Titles

Figure 1 shows the increase in anonymous hiring during the pilot period in the treated job titles. The take-up rate was lower in the first months of the pilot when only randomly selected job openings were subject to anonymization.⁹ Based on our data, 44 of 99 or

⁹City of Helsinki decided to stop randomizing because lower than expected number of job openings becoming assigned to the treatment group.

44 percent of the recruitment processes were verifiably 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 as it was done manually by HR. However, we cannot exclude the possibility that some of this might be driven by hiring managers' decisions. It is also possible that the knowledge of being in the treatment group can also affect hiring independent of anonymization. Thus, we will always compare differences between the treatment and control groups as a whole as defined prior to the pilot. Out of the treated processes 62 resulted in hiring someone.

5.2 The Effects on Hiring Probabilities

Table 3 shows our main results. First, in columns (1) to (3) in Panel A, we find that applicants with a foreign-sounding name are more likely to be called to an interview when applying for a job that is part of the anonymous hiring pilot. The point estimate on the baseline specification with only job title fixed effects suggests that interview probability increases by 5 percentage points relative to applicants with Finnish-sounding names. The estimate changes little if we add individual-level controls for demographics (Column 2) or skills (Column 3) which could be otherwise important determinants of hiring decisions. Overall applicants with a foreign-sounding name receive a call for an interview with a probability of 4 percent while the same number for applicants with a Finnish-sounding name is 8 percent. This suggests that the effect of anonymization is about equal in magnitude to the gap in interviews between the two groups. Second, in Columns (4) to (6) we show results on hiring decisions where we see a positive effect which is similar or somewhat larger in magnitude. Again, we do not observe any meaningful changes in the estimates when we control for individual-level characteristics. These results suggest that anonymous hiring not only affects the early stage of the hiring process but that it carries over to the final hiring decision.

In our baseline results, we use first names as the proxy for ethnic background. To show how sensitive our results are to the measure of ethnic background, we first look at hires by the applicants' first language. We use non-native first language as a measure of foreign ethnicity which has the benefit that it is a measure that is independent of our baseline measure but highly correlated.¹⁰ Table 4 Panel A shows that the effects are qual-

¹⁰See the appendix for correlations.

itatively similar to our baseline specification. The point estimates for calls to interviews are somewhat smaller and statistically insignificant at traditional levels, but the effects on hiring decisions remain statistically significant and similar in magnitude. Second, one could argue that different types of names carry different types of information about the applicant that a manager might care about. For example, a foreign last name or a foreign first name combined with a Finnish first name and last name could send different signals to managers.¹¹ To study this, we show in Panels B to D that the effects are similar if we use information on foreign-sounding full names, any name, or last name only. Again, we find qualitatively similar results. The estimates go in the expected direction for the measure based on last names as that is where we are expected to classify what are actually Finnish names to foreign names more often and thus have more measurement error.

While the main motivation for the pilot was to improve the hiring outcomes of applicants with foreign names or with immigrant backgrounds, it also provided an opportunity to study the effects of other types of potential discrimination. More specifically, the removal of applicants' names made the policy of not eliciting applicants' gender much more stringent. However, in contrast to the effects found for changes in the ethnic decomposition of hires, we do not observe statistically significant changes in the gender composition. Panel B shows the estimates on the probability of women being called to interviews or hired. The point estimates for interviews are negative but statistically insignificant and the point estimates for hiring decisions are very close to zero. It needs to be noted that on average for the hiring decisions that we consider, women have a higher probability of being called to an interview or to be hired than men. However, our estimates are not precise enough that we could completely rule out potentially meaningful impacts in either positive or negative direction.

Finally, we also conduct an event-study analysis of our main results. Figure 2 shows the differences in the probability of being interviewed and hired for applicants with foreign-sounding names vs. Finnish-sounding names (Panels A and C) and for women vs. men (Panels B and D) in the treated vs. control job titles around the introduction of the pilot program. For the analysis, we divide the sample period into 10-week periods to have two pre-treatment periods which allows us to observe trends before the pilot period but also to have enough job postings in each period to have more statistical power.

¹¹In the appendix, we show that all these measures are highly correlated as expected. The correlation between first names and last names is the lowest. This is as expected as there is likely the most measurement error on the last name measure. Many uncommon last names easily become classified as foreign even if they sound Finnish.

First, we see that prior to the introduction of the pilot, there appears to be no differential trend between the control and treatment groups. This provides some evidence that our assumption of common trends in these groups remains plausible. Second, the results remain consistent with our baseline difference-in-difference results. In particular, we see that there is an increase in interviews and hiring for applicants with foreign-sounding names. The estimates in Panel C suggest that the probability increases by more than 5 percentage points relative to applicants with Finnish-sounding names and remains about constant during the pilot period. Panel A shows a more gradual increase for interviews with point estimates around the 5 percentage point mark. Similarly consistent with our baseline results, we see no statistically significant changes in interviews or hiring by gender.

5.3 The Effects on Applications and Total Hiring

Our setting and data allow us to also measure the effects of the pilot on application choices which could be important in driving the total effects of the pilot. 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 and, as a consequence, to hire more people from these groups.

Tables 5 and 6 show our main results on applications. First, our results suggest that job titles that were anonymized and also advertized as anonymous received more applications (Table 6, Column 1). However, it does not appear that this increase is driven mainly by individuals with foreign-sounding applicants who were directly targeted with the policy and seem to have had the most to gain from the policy. In particular, we find that if anything, there appears to be a reduction in the share of applications from individuals with foreign-sounding names (Table 5, Panel A, Column 1). The point estimate suggests that the share goes down by 3 percentage points, which is not statistically significant, but potentially meaningful in size relative to the baseline share in applications. On the other hand, we observe that the share of women increases when hiring processes are advertised as anonymous (Table 5, Panel B, Column 1). The point estimate is about 6 percentage points. Taken together, it seems that advertising anonymous hiring can affect application decisions but not always in the expected way.

The total effects of the anonymization policy are a combination of changes in hiring probability among applicants as well as changes in the applicant composition. Thus, to measure the total effects that combine these two channels, Table 5 columns (2) and

(3) show how the composition of interviewees and hirees changes in the anonymized processes vs. the standard processes. Our results suggest that the combination of i) a reduction in the share of applicants and ii) an increase in the interview and hiring probability results in a positive effect on total interviews and hiring for applicants with foreign-sounding names. These estimates are now less precise in part due to the fact that we may no longer take advantage of the individual level variation and compare outcomes at the job opening level but also due to the fact that 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 in total there are actually more women interviewed and hired under the anonymized processes. However, these estimates are also not statistically significant at standard levels. These results however illustrate the two potential mechanisms through which anonymization may affect final hiring outcomes which employers may care about.

We then repeat the same type of analysis at the applicant level. The benefit of the individual-level analysis is that it allows us to control for applicants' other characteristics that could also change due to the policy. For example, more skilled individuals might apply for anonymous jobs. Comparing Appendix Table A1 columns (1) and (2) shows that individual-level controls do not change these results so the effect does not seem to be driven by simply more skilled workers applying.

The results suggest that job applicants respond to the announcement of an anonymized hiring process. It appears that women may perceive their chances to improve or that the anonymization sends a positive signal about the job or potential colleagues or the manager. In contrast, for applicants with a foreign-sounding name, the effect is, if anything, reversed. It could be that announced anonymization sends a negative signal about the job attributes, for example, a lack of previous ethnic hires. However, our setting does not allow us to directly infer the source of these choices.

Table 6 shows how anonymization affects the number of interviews and the probability that someone was hired. First, we find that on average there was no change in the number of job interviews. This finding does not support the idea that hiring managers would opt to interview more job candidates to have more discretion after interviews and after observing job applicants' full background characteristics. Second, we do not see any changes in the share of job openings that result in hiring someone. Thus we do not find evidence that it would have been harder to find suitable candidates. However, there is a caveat in our data concerning the latter. In particular, in our data, we do not always

observe the individual that was ultimately hired even if the job opening was closed and we observe that someone was hired to the open position. In fact, we observe that during the pilot, this information is missing more often 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 and thus this is a non-negligent 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 the HR and the hiring managers did not always use the recruitment system to keep track of the final hiring outcomes.

5.4 The Effect on Job Performance and Wages

To understand whether individuals are more or less suitable for the jobs that they are hired for, we next compare the post-hiring outcomes of workers hired in the treatment vs. control groups. Figure 3 shows our main results. Panel A shows the share of job applicants working at the city of Helsinki around the time of the job start. First, it shows that on average more than 50 percent of workers hired have already an existing contract with the City of Helsinki before they start on the new job. After 10 months of job start, about 80 percent are still working at the City of Helsinki. Now, Panel B shows that there are no significant differences in employment between workers hired in anonymous vs. standard processes in the first ten months after the job starts. The point estimates are close to zero, but the precision of the estimates decreases over time.

Next, we may look at the earnings of workers in Panel C and D. On average, workers earn about 2,700 euros in the month of job start and this increases somewhat in the first two months but then starts to decline as more and more of the workers are leaving the City of Helsinki. However, again, there exist no statistically significant differences in the earnings of individuals hired anonymously relative to the standard hiring processes. The point estimates are initially positive but quite small and close to zero.

We can also estimate a differences-in-differences effect for the total wage and months worked during the first 11 months after recruitment. This should give us more statistical power, but again we find no significant effect of anonymization (Table 7). The point estimates suggest that during the first 11 months applicants who were hired to job titles with the anonymization policy earned about 220 euros (or one percent) more and worked for 0.03 months (0.3 percent) longer than those in the control group. The point estimates are thus very small, but the estimates remain quite imprecise and we cannot confidently rule out economically significant effects in either direction. In Table 7 Column (3) we also show

that anonymous processes did not differ in terms of how quickly work started counting from the end of the recruitment process. The point estimate suggests that on average the time to job start decreased by 6.6 days. Thus, even if the anonymization meant that hiring managers could not start calling applicants to interviews immediately after applications were received, this does not appear to significantly delay the recruitment and job-starting process.

Taken together, we find no evidence that workers hired via an anonymous recruitment process would be less suitable for the jobs than under standard hiring processes. However, the precision of the estimates is still quite low, so we cannot rule out either potentially meaningful positive or negative effects on employment or earnings.

6 Hiring Managers' Selection into Anonymizing

Previous research has shown evidence of unintended negative effects of anonymous hiring processes on ethnic or minority hiring decisions. The suggested explanation is that if firms or managers can opt in to anonymization, it may be that self-selection induces firms or managers with discriminatory preferences or hiring styles not to participate. If managers and firms, who would otherwise positively discriminate, self-select into anonymization, there would be an aggregate deterioration in the position of the discriminated groups.¹² This suggests that the way the policy is implemented may matter for how effective it is likely to be.

Our empirical setting allows us to study whether self-selection could explain the difference between our main results and previous empirical evidence. During the pilot, managers do not individually select to participate, but the decision is driven by the job titles selected for the pilot. This should exclude the possibility of the mechanism suggested in previous research. However, in early 2021 and after the pilot period ended, the City of Helsinki made anonymization an opt-in policy. Hiring managers could voluntarily run the hiring process anonymously but were not mandated to do so. Thus, we can use this policy change to study the differences in the managers who decide to anonymize when it becomes voluntary.

Table 8 describes selection into anonymization by manager and job characteristics during and after the pilot when it is made an opt-in policy. We find that manager-level char-

¹²Positive discrimination is only one possible mechanism. It could also be that some managers interpret the job experience of minorities in a different way than for others.

acteristics or job attributes are not predictive of the choice to anonymize the hiring process during the pilot stage (Columns 1 to 3). In particular, we cannot reject that all coefficients are equal to zero. This result remains if we control for the division or broad occupation-level fixed effects. This is consistent with the fact that anonymization was not an individual manager's choice but determined at a higher level at the City of Helsinki. However, once the anonymization policy is changed after the pilot, we see that there emerges an important element of selection into anonymization. Managers who opt in are less likely to be women, more likely to be younger, and Finnish speaking. Now, we do not have any direct individual-level measure of discriminatory tendency that would allow us to say if those who participate are more or less likely to discriminate. Thus it is hard to predict what the likely direction this selection has on hiring outcomes. However, these results suggest that hiring managers who decide to use anonymization are different than those who do not and it is not clear that the effects of a voluntary anonymization policy are going to be similarly effective as the mandated anonymization policy.

7 Conclusion

In this paper, we leverage the City of Helsinki's anonymous hiring pilot to provide new evidence on the effects of anonymous hiring. Our results suggest that anonymous hiring may serve as a potentially effective tool to level the playing field between job applicants with foreign-sounding names and those with Finnish-sounding names. Our setting allows us to shed new light on the potential mechanisms through which this effect may come about. We may show that application decisions may be important and we may use data on post-hiring employment duration and wages as proxies for match quality. We can also leverage a change in the anonymization policy to study how the design in the implementation (mandated or voluntary) may affect the types of managers who decide to take up anonymization. Yet, limitations in the pilot sample size mean that for some of these outcomes, we cannot rule out potentially meaningful effects. This highlights the need for more research on these questions.

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.

- Ahmad, A. (2020). When the name matters: An experimental investigation of ethnic discrimination in the Finnish labor market. *Sociological Inquiry*, 90:468–496.
- Åslund, O. and Skans, O. N. (2012). Do Anonymous Job Application Procedures Level the Playing Field? *ILR Review*, 65(1):82–107.
- Behaghel, L., Crépon, B., and Le Barbanchon, T. (2015). Unintended Effects of Anonymous Résumés. *American Economic Journal: Applied Economics*, 7(3):1–27.
- 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.
- BETA (2017). Going blind to see more clearly: Unconscious bias in Australian Public Service shortlisting processes. Technical report.
- Doleac, J. L. and Hansen, B. (2018). The unintended consequences of ‘ban the box’: Statistical discrimination and employment outcomes when criminal histories are hidden. page 63.
- Goldin, C. and Rouse, C. (2000). Orchestrating Impartiality: The Impact of "Blind" Auditions on Female Musicians. *The American Economic Review*, 90(4):715–741.
- Hoffman, M., Kahn, L. B., and Li, D. (2018). Discretion in Hiring*. *The Quarterly Journal of Economics*, 133(2):765–800.
- Kessler, J., Low, C., and Sullivan, C. (2019). Incentivized Resume Rating: Eliciting Employer Preferences without Deception. Technical Report w25800, National Bureau of Economic Research, Cambridge, MA.
- 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. (2012a). Anonymous job applications in Europe. *IZA Journal of European Labor Studies*, 1(1):5.
- Krause, A., Rinne, U., and Zimmermann, K. F. (2012b). Anonymous job applications of fresh Ph.D. economists. *Economics Letters*, 117(2):441–444.

- 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.
- Pager, D. (2003). The Mark of a Criminal Record. *American Journal of Sociology*, 108(5):937–975.
- Public Service Commission of Canada (2018). Anonymized Recruitment Pilot Project - Final report. <https://www.canada.ca/en/public-service-commission/services/publications/Name-blind-recruitment-pilot-project.html>.
- Quillian, L., Pager, D., Hexel, O., and Midtbøen, A. H. (2017). Meta-analysis of field experiments shows no change in racial discrimination in hiring over time. *Proceedings of the National Academy of Sciences*, 114(41):10870–10875.
- Veuger, S. and Shoag, D. (2019). 'Ban the Box' measures help high-crime neighborhoods. AEI Economics Working Paper, American Enterprise Institute.

Figures and Tables

Table 1: Applicants, Jobs, and Hires in the Pilot

	Applicants		Vacancies		Hires	
	Pre-pilot	Pilot	Pre-pilot	Pilot	Pre-pilot	Pilot
Treated	930	2,773	49	99	40	62
Control	25,971	63,568	2,440	5,494	1,392	3,090
Total	26,901	66,341	2,489	5,593	1,432	3,152

Notes: This table shows the number of applicants, jobs, and hired individuals in job titles during the pilot period from September 2019 to December 2020. The pre-period is from September 2019 to January 2020. The pilot-period goes from February 2020 to December 2020.

Table 2: Pre-Treatment Characteristics of Job Openings, Applicants, and Recruits

	All (1)	Control (2)	Treated (3)	(3)-(2) (4)	S.E. (5)	N (6)
Panel A: Job Characteristics						
No. Applicants	32.27	31.97	43.73	11.76***	(1.10)	18,273
Temporary Contract	0.23	0.23	0.19	-0.04**	(0.02)	18,273
Panel B: Applicant Characteristics						
Woman	0.79	0.79	0.74	-0.05**	(0.02)	18,273
Foreign Sounding First Name	0.07	0.07	0.11	0.04***	(0.01)	18,273
Finnish Native Language	0.93	0.93	0.87	-0.06***	(0.02)	18,273
Age	37.86	37.83	39.11	1.28**	(0.56)	18,273
Bachelor's Degree'	0.66	0.67	0.28	-0.40***	(0.02)	18,273
Experience (Years)	10.05	10.04	10.50	0.46	(0.45)	18,273
Interview	0.09	0.09	0.06	-0.02**	(0.01)	18,273
Panel C: Recruit Characteristics						
Woman	0.80	0.81	0.59	-0.22*	(0.12)	945
Foreign Sounding First Name	0.07	0.07	0.00	-0.07***	(0.01)	945
Finnish Native Language	0.95	0.95	1.00	0.05***	(0.01)	945
Age	38.30	38.35	35.30	-3.05	(2.10)	945
Bachelor's Degree'	0.62	0.62	0.35	-0.27**	(0.12)	945
Experience (Years)	10.54	10.54	10.30	-0.25	(2.09)	945

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, Columns (2) and (3) show the characteristics separately for job openings in the treatment and control groups. Column (4) shows the difference in control and treatment group averages. Robust standard errors in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table 3: The Effect of Anonymization on Hiring and Calls to Interviews

	P(Interview)			P(Hired)		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Foreign Sounding First Name						
Foreign X Treat X Post	0.049* (0.029)	0.052* (0.029)	0.049* (0.029)	0.072*** (0.015)	0.073*** (0.015)	0.071*** (0.015)
N	74,803	74,803	74,803	74,803	74,803	74,803
Foreign Mean	0.040	0.040	0.040	0.007	0.007	0.007
Finnish Mean	0.083	0.083	0.083	0.050	0.050	0.050
Job FE	✓	✓	✓	✓	✓	✓
Demographic controls		✓	✓		✓	✓
Skill controls			✓			✓
Panel B: Women						
Woman X Treat X Post	-0.022 (0.033)	-0.023 (0.032)	-0.023 (0.032)	0.001 (0.022)	0.001 (0.022)	0.001 (0.022)
N	74,803	74,803	74,803	74,803	74,803	74,803
Women Mean	0.086	0.086	0.086	0.042	0.042	0.042
Men Mean	0.059	0.059	0.059	0.054	0.054	0.054
Job FE	✓	✓	✓	✓	✓	✓
Demographic controls		✓	✓		✓	✓
Skill controls			✓			✓

Notes: This table shows difference-in-difference estimates using the specification REF. Columns (1) and (4) control for job title fixed effects and provide the baseline estimates. Column (2) and (5) and controls for applicants age and gender. Columns (3) and (6) add controls for applicants' job experience and educational attainment. 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 4: The Effect of Anonymization on Hiring and Calls to Interviews Using Alternative Measures of Ethnic Identity

	P(Interview)			P(Hired)		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Foreign Native Language						
Foreign X Treat X Post	0.019 (0.021)	0.023 (0.021)	0.023 (0.020)	0.041*** (0.012)	0.043*** (0.012)	0.043*** (0.012)
Foreign Mean	0.026	0.026	0.026	0.007	0.007	0.007
Finnish Mean	0.086	0.086	0.086	0.050	0.050	0.050
Panel B: Foreign Sounding Full Name						
Foreign X Treat X Post	0.028 (0.029)	0.032 (0.029)	0.031 (0.029)	0.044*** (0.013)	0.047*** (0.013)	0.045*** (0.013)
Foreign Mean	0.036	0.036	0.036	0.000	0.000	0.000
Finnish Mean	0.082	0.082	0.082	0.049	0.049	0.049
Panel C: Any Foreign Name						
Foreign X Treat X Post	0.044* (0.023)	0.047** (0.023)	0.046** (0.023)	0.046*** (0.015)	0.048*** (0.015)	0.047*** (0.015)
Foreign Mean	0.048	0.048	0.048	0.020	0.020	0.020
Finnish Mean	0.087	0.087	0.087	0.052	0.052	0.052
Panel D: Foreign Last Name						
Foreign X Treat X Post	0.031 (0.022)	0.035 (0.022)	0.035 (0.022)	0.026* (0.015)	0.028* (0.015)	0.028* (0.015)
Foreign Mean	0.047	0.047	0.047	0.019	0.019	0.019
Finnish Mean	0.085	0.085	0.085	0.050	0.050	0.050
N	74,803	74,803	74,803	74,803	74,803	74,803
Job FE	✓	✓	✓	✓	✓	✓
Demographic controls		✓	✓		✓	✓
Skill controls			✓			✓

Notes: This table shows difference-in-difference estimates using the specification REF. Columns (1) and (4) control for job title fixed effects and provide the baseline estimates. Column (2) and (5) and controls for applicants age and gender. Columns (3) and (6) add controls for applicants' job experience and educational attainment. 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 5: The Effects of Anonymization on Applications and Total Hiring at the Job Level

	(1) Applicants	(2) Interviewees	(3) Hired
Panel A: Foreign Sounding First Name			
Treat X Post	-0.028 (0.020)	0.081* (0.048)	0.077 (0.056)
<i>N</i>	3,714	3,714	3,714
Outcome Mean	0.125	0.030	0.025
Job FE	✓	✓	✓
Panel B: Woman			
Treat X Post	0.059* (0.035)	0.058 (0.095)	0.111 (0.099)
<i>N</i>	3,714	3,714	3,714
Outcome Mean	0.721	0.740	0.700
Job FE	✓	✓	✓

Notes: This table shows difference-in-difference estimates using the specification similar to REF aggregated to the job opening level. Robust standard errors in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. The outcome means are pre-treatment period means for the treated group.

Table 6: The Effects on the Number of Applications, Interviews and Hires

	(1) No. Applicants	(2) No. Interviewed	(3) Someone Hired
Treat X Post	5.695** (2.665)	0.169 (0.287)	0.055 (0.061)
<i>N</i>	7516	7516	7516
Outcome Mean	18.980	1.449	0.837
Job FE	✓	✓	✓

Notes: This table shows difference-in-difference estimates using the specification similar to REF aggregated to the job opening level. The outcome mean is the mean for the treated group in the pre-period. Robust standard errors in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table 7: Differences-in-Differences Estimation of the Total Wage, Months Worked, and Days Until Job Start for the First 10 Months on the Job

	Wage	Months Worked	Days Until Job Start
Treat X Post	218.987 (1425.670)	0.031 (0.469)	-6.646 (4.762)
<i>N</i>	2848	2848	2848
Outcome Mean	24274.001	10.111	45.107
Job FE	✓	✓	✓

Notes: This table shows the difference-in-differences estimates for the hired applicants in the first 10 months of their employment at the City. The outcome mean is the mean for the treated group in the pre-period. Robust standard errors in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

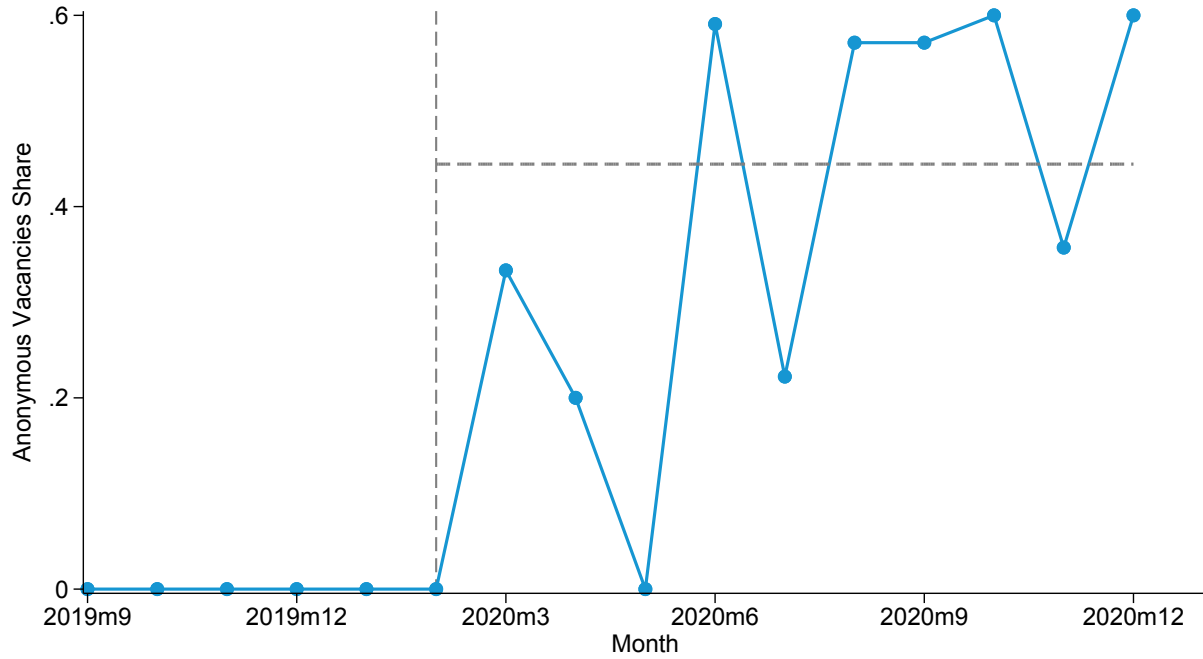
Table 8: Recruiters' Self-selection into Anonymization (in Percentage Points)

	Pilot			Post Pilot		
	(1)	(2)	(3)	(4)	(5)	(6)
Woman Recruiter	-0.143 (0.176)	-0.015 (0.195)	0.068 (0.177)	-2.433** (1.103)	-1.726 (1.106)	-2.046* (1.144)
Recruiter's Wage (1000)	-0.208 (0.149)	-0.233 (0.163)	-0.362* (0.212)	0.146 (0.171)	-0.056 (0.116)	-0.055 (0.145)
Recruiter's Age	0.019 (0.019)	0.019 (0.018)	0.023 (0.017)	-0.072*** (0.024)	-0.057*** (0.020)	-0.062*** (0.021)
Native Language Finnish	0.390* (0.226)	0.324 (0.219)	0.271 (0.211)	1.557*** (0.407)	1.071*** (0.393)	1.317*** (0.437)
No. Vacancies	-0.008 (0.006)	-0.007 (0.004)	-0.010 (0.007)	-0.005 (0.008)	-0.008 (0.011)	-0.008 (0.011)
Temporary Contract	-0.112 (0.209)	-0.090 (0.200)	-0.167 (0.208)	-0.648* (0.367)	-0.630* (0.318)	-0.548 (0.422)
Outcome Mean	0.426	0.426	0.426	1.619	1.619	1.619
<i>N</i>	6,099	6,099	6,099	9,946	9,946	9,946
Division FE		✓	✓		✓	✓
Occupation FE			✓			✓
F-statistic	0.764	0.742	1.142	4.181	1.844	2.030
P-value	0.600	0.617	0.345	0.001	0.099	0.070

Notes: This table shows descriptive evidence of managers' self-selection into anonymized recruitment processes. Columns (1) to (3) show how recruiter and job characteristics predict anonymization during the pilot period when individual managers are not allowed to choose whether to anonymize or not. Columns (4) to (6) show the same results after the pilot when anonymization is made voluntary for recruiters. Point estimates are expressed in percentage points. Robust standard errors in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

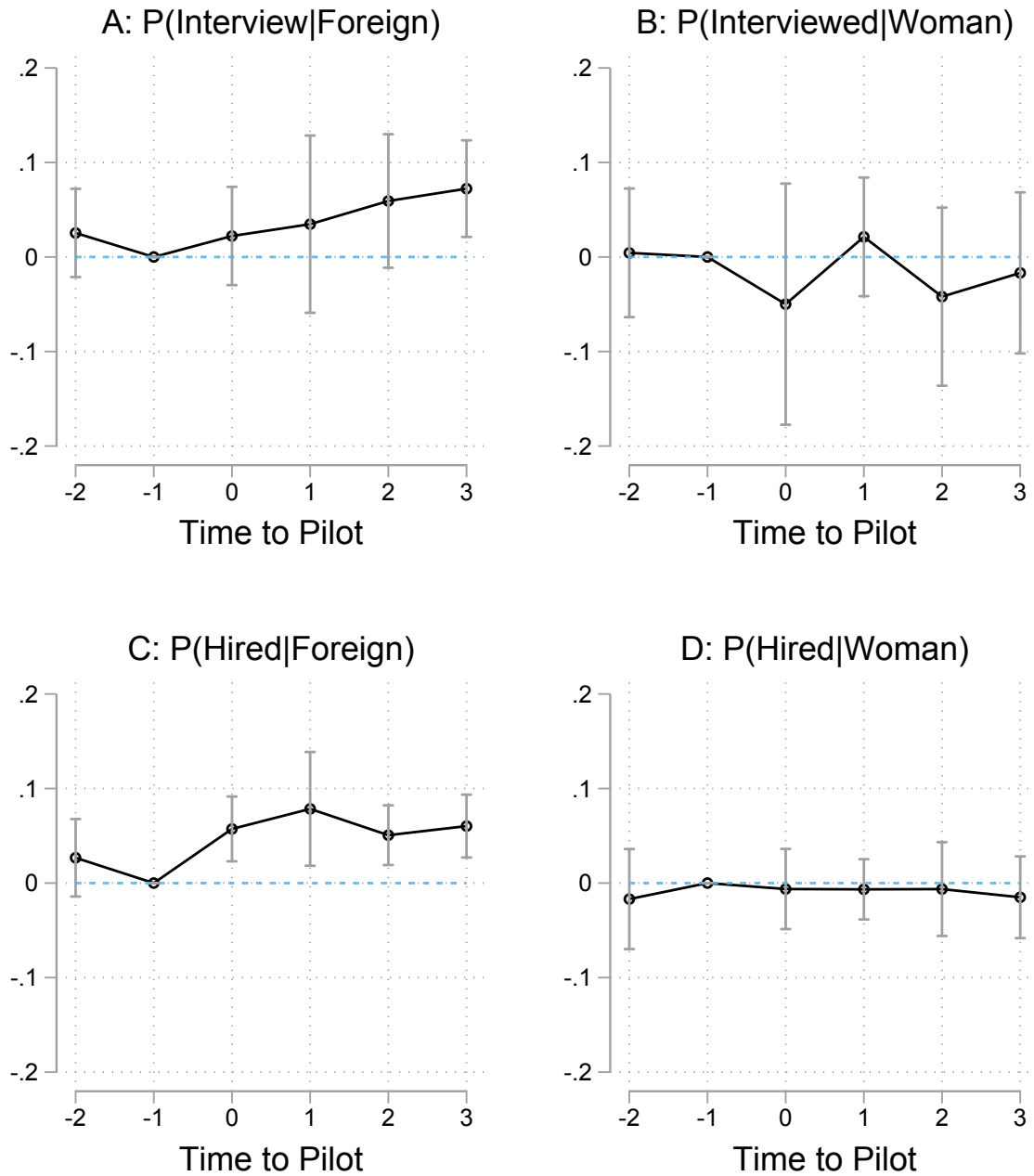
Figures

Figure 1: Anonymization in Treated Job Titles



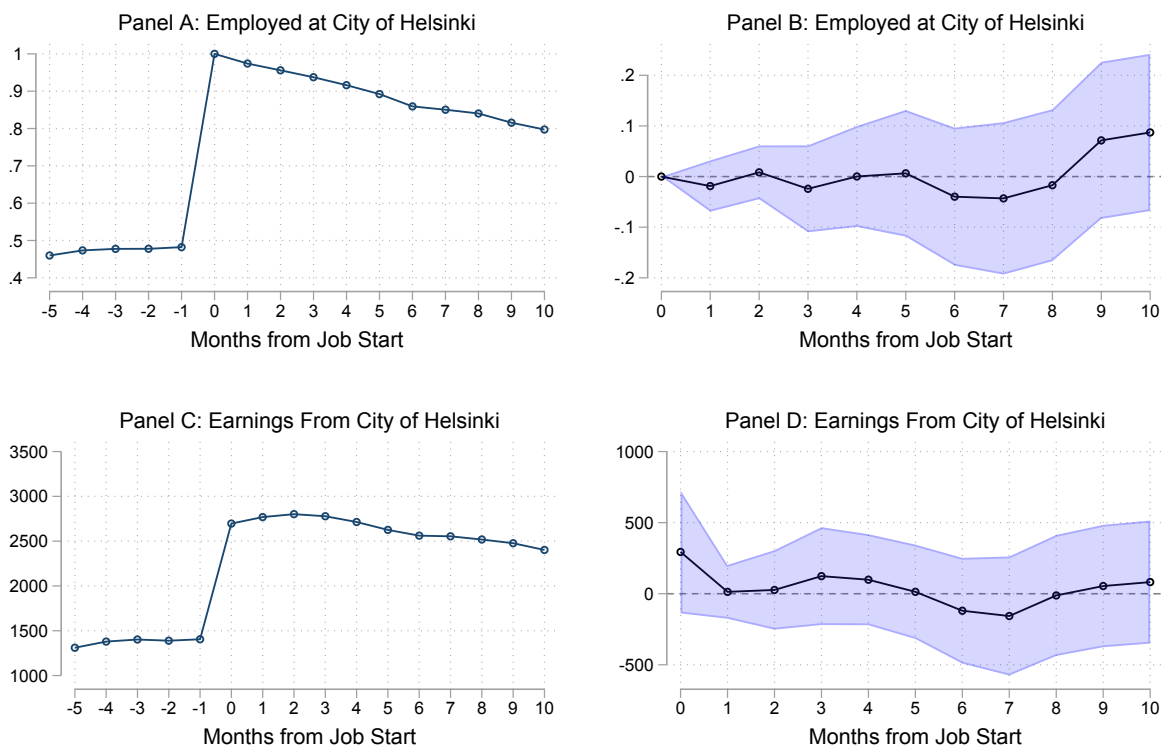
Notes: This figure shows the share of anonymized recruitment processes at the City of Helsinki around the pilot period in the treatment group. The dashed horizontal line marks the beginning of the pilot program. The dashed vertical line shows the average during the whole pilot period.

Figure 2: Event Study Estimates on Hiring Outcomes



Notes: This figure shows the event study estimates on the hiring of applicants with foreign-sounding names compared to applicants with Finnish-sounding names and women compared to men. The time periods have been aggregated 10-week periods. Standard errors are clustered at the job opening level. The figure shows 95% confidence intervals.

Figure 3: 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

Appendix

Table A1: The Effects of Anonymization on Applications and Total Hiring at the Individual Level

	Applicants		Interviewees		Hired	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Foreign Sounding First Name						
Treat X Post	-0.013 (0.014)	-0.009 (0.015)	0.058 (0.048)	0.061 (0.048)	0.147** (0.061)	0.162*** (0.062)
<i>N</i>	53,134	53,134	5,561	5,561	3,575	3,575
Outcome Mean	0.115	0.115	0.059	0.059	0.000	0.000
Job FE	✓	✓	✓	✓	✓	✓
Demographic controls		✓		✓		✓
Skill controls		✓		✓		✓
Panel B: Woman						
Treat X Post	0.066*** (0.021)	0.064*** (0.022)	-0.049 (0.091)	-0.040 (0.090)	0.127 (0.116)	0.132 (0.117)
<i>N</i>	53,134	53,134	5,561	5,561	3,575	3,575
Outcome Mean	0.750	0.750	0.804	0.804	0.679	0.679
Job FE	✓	✓	✓	✓	✓	✓
Demographic controls		✓		✓		✓
Skill controls		✓		✓		✓

Notes: This table shows difference-in-difference estimates using the specification similar to REF at the individual level allowing us to use individual-level controls. Columns (1), (3), and (5) control for job title fixed effects and provide the baseline estimates. Columns (2), (4), and (6) control for applicants' age and gender as well as job experience and educational attainment. Robust standard errors in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. The outcome means are pre-treatment period means for the outcome variables.

Table A2: Correlations Between the Foreign Variables

	First Name	Last Name	Any Name	Full Name	Language
First Name	1				
Last Name	0.519***	1			
Any Name	0.699***	0.902***	1		
Full Name	0.845***	0.655***	0.590***	1	
Language	0.616***	0.598***	0.591***	0.665***	1
Mean	0.152	0.229	0.268	0.113	0.137

Notes: This table shows correlations between different measures of foreign applicants. $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

A Linking Hires and Recruiters 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 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.

To link these two data, we use the information on the names and birth dates 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 both data sets. We transform the names into LATIN-ASCII characters to remove any special characters. We use the *relink* package in Stata which creates a score between 0 and 1 for the matches. The value of 1 is a perfect match. In the hiree data, we have 10,566 unique applicant IDs. Some applicants have two different applicant IDs. This could happen e.g. if they created a new user with a different email address. We remove these applicants (288).

We found 8,579 matches in total and out of those, 8,069 perfect matches. After assessing the imperfect matches, we deleted all matches with a score lower than .7 (426). Matches with a lower than .7 score are likely to be false based on visually observing the data.

This leads us to 8,162 hires whom we can find in the payroll data. Thus, we match 83

percent of hires to an employee in the payroll data (8,579/10,278).

Next, we need to match all hires to the correct job spells or contracts at the city of Helsinki. This is necessary as many of the hires are internal and some may apply and be hired to multiple jobs during our sample period. To do so, we match each applicant to their first job spell that starts after the vacancy has been closed.

Finally, if an applicant has multiple job spells in one month, we keep the job spell that started closest to the application period's end. We restrict the time between the application period ending and a job spell starting to six months. This cut our data to 7,995 job spells (79%). Individuals in the data had 0 to 4 jobs in one month. We use these job spells in our analysis. We form a balanced panel by filling in all the months in our analysis (from September 2019 to February 2022).

For the recruiters in each recruitment, we only have their (self-reported) names. These names can be (and in most cases are) different from the names they are under in the payroll data. We use *relink* to match recruiters to the payroll data. The payroll data gives us an idea of the recruiters' wage, age, and native language. We are able to match 72% recruiters to a worker in the payroll data.

Table A3: Forming a Panel Data by Linking the Application Data to the Payroll Data

Data	N	Match Rate
Hired Applicants	10,566	
Employees	76,715	
Hiree matched to an employee in the payroll	8,579	83 %
Hirees matched to a job spell	7,995	79 %
Applicant data anonymously hired applicants	76	
Anonymously hired job spells	57	
Recruiters in Vacancies	2,378	
Recruiters matched	1,706	72%



Työn ja talouden tutkimus Labore on vuonna 1971 perustettu itsenäinen ja voittoa tavoittelematon kansantalouden asiantuntijayksikkö. Laitoksessa tehdään taloustieteellistä tutkimusta ja laaditaan suhdanne-ennusteita. Lisäksi laitoksen tutkijat toimivatulkopuolisissa asiantuntijatehtävissä sekä osallistuvat aktiivisesti julkiseen talouspoliittiseen keskusteluun. Laboren toiminnan tavoitteena on tarjota tutkimustietoa yhteiskunnallisen keskustelun sekä päätöksenteon tueksi.

Laboressa tehtävän tutkimustyön painopiste on tilastollisiin aineistoihin perustuvassa empiirisessä tutkimuksessa. Sen taustalla on vahva teoreettinen näkemys ja tieteellisten menetelmien asiantuntemus.

Työn ja talouden tutkimus Labore
Arkadiankatu 7 (Economicum)
00100 Helsinki
Puh. +358 40 940 1940
labore.fi

ISBN 978-952-209-201-4
ISSN 1795-1801 (pdf)