The Effects of Gender and Parental Occupation in the Apprenticeship Market: An Experimental Evaluation

Ana Fernandes* Martin Huber† Camila Plaza‡

WORK IN PROGRESS

March 5, 2019

Abstract

This study investigates the effects of applicant gender and parental occupation on callback rates in the Swiss apprenticeship market, i.e. invitations to an interview or trial apprenticeship. It sheds light on the question of whether employer behaviour in terms of callbacks contributes to occupational segregation, i.e. differences in occupational choices across gender and socio-economic status, when considering the earliest point of entry into the labor market. We ran a correspondence test and sent out fictitious job applications with randomized gender and parental occupation to apprenticeship vacancies in several Swiss regions. We generally find no robust evidence for a differential treatment by employers, as gender and parental occupation do not statistically significantly affect callback rates in most cases. The one exception is stating father's occupation to be university professor, which statistically significantly boots call backs for female applications even when accounting for multiple hypothesis testing, but not for male applications. This suggests that applications should ideally not reveal socio-economic information.

JEL Classification: C93, J16, J71

Keywords: Experimental economics, Discrimination, Parental Occupation, Gender

We are indebted to André Scholl for the technical implementation and management of the correspondence test and to Aron Baeriswyl for his support in the conceptual preparation and documentation of the project. We thank either of them as well as Benjamin Bolzern and Bénédicte Droz for their work and effort during the application process and are also grateful to Ruth Neuhaus and Andrea Sommer-Gauch for their help and administrative support. Financial support from the Swiss National Science Foundation for the SNF project 100018_176376 'Gender Occupational Segregation in the Swiss Apprenticeship Market: the Role of Employers in an Experimental Evaluation' is gratefully acknowledged.

*Bern University of Applied Sciences, Brückenstr. 73, CH-3005 Bern, ana.fernandes@bfh.ch

†University of Fribourg, Bd. de Pérolles 90, CH-1700 Fribourg, martin.huber@unifr.ch

‡University of Basel, Peter Merian-Weg 6, CH-4002 Basel, camila.plaza@unibas.ch

1 Introduction

In this paper, we experimentally assess how applicant gender and parental occupation affect call back rates in the Swiss apprenticeship market, namely invitations to an interview or trial apprenticeship when applying to vacancies. Our study is motivated by the stylized fact that occupational choices generally differ across gender and socio-economic status. Specifically, we are interested in whether employer behaviour contributes to such occupational segregation at the earliest (legally) possible entry into the labor market by a differential treatment of applications across gender and/or parental occupation.

Occupational segregation w.r.t. gender is but one of numerous manifestations of the asymmetric way in which men and women participate in the labor market (see e.g. Blau and Kahn (2017) or Cortes and Pan (2018)). While there is little disagreement over the facts, there is considerable less agreement on the causes of those empirical regularities. One of our goals is to study whether or not employers contribute to gender occupational segregation by more likely rejecting a specific gender, in particular if it is usually not associated with the corresponding occupation (e.g., females are under-represented in technical fields). The work of Huggett et al. (2011) has shown that early shocks in a person's labor market history tend to be magnified in future outcomes. We therefore conduct our study on the first point of entry into the labor market in developed economics, the apprenticeship market.

In Switzerland, job applications routinely contain personal information, including a picture and personal details (such as age, marital status, and so on). Apprenticeship applicants are usually 14 or 15 years of age. Because of their youth, they usually do not yet have that much to say about themselves in their CVs. However, they often indicate the profession of their parents. This unique feature of the Swiss apprenticeship market allows us to further investigate whether parental background affects the labor market chances of offspring. This is an important question as equality of opportunity would require such background information not to have an effect on the applicant's labor market outcomes. How closely one's earnings relate to those of one's parents is the subject of an extensive literature attempting to estimate the *intergenerational elasticity* (IGE), a measure of intergenerational income persistence. To the best of our knowledge, this is the first attempt to address the intergenerational persistence of income using an experimental data collection method.

In economics, asymmetric labor market treatment of job applicants for reasons unrelated to their ability to perform the job amounts to *discrimination*, see Becker (1957). The two main reasons for employers to discriminate offered in the literature originate in tastes, e.g. employers or customers dislike working with a particular group in the population, or uncertainty about the true productivity of the candidate employee, see Arrow (1971) and Phelps (1972). The former is commonly known as *taste-based discrimination* whereas the latter amounts to *statistical discrimination*. The preference for one gender over the other in relation to the occupation sex type could have elements of both taste-based and statistical discrimination. Employers may have a preference for candidates having the gender that matches the sex type of the occupation, possibly reflecting stereotypical preference biases; they may also believe that matching the sex type of the occupation is relevant for productivity. See Weichselbaumer (2004) for a detailed discussion on this matter. An interesting aspect is that due to the young age of apprenticeship applicants, statistical discrimination against females due to family or fertility planning appears less likely in our context than for older age groups.

In order to assess whether employers take applicant gender and parental occupation into consideration when evaluating applications for apprenticeships, we conducted a *correspondence testing* experiment. We to this end sent out 2940 fictitious applications containing CVs and educational certificates to open apprenticeship positions across four regions in Switzerland, namely Basel, Bern, Lausanne, and Zurich, between August and October 2018. In the applications, we randomized demographic characteristics like gender and parental occupation to investigate the impact on call back rates by employers, namely invitations to interviews or trial apprenticeships up to February 2019. Using applications that (at least on average) signalled a comparable level of productivity and varied only w.r.t. applicant gender and/or parental occupation, was key for investigating whether employers systematically differ in their treatment of groups with particular demographics.

Concerning parental occupation, the CV either stated that the mother worked as a primary school teacher or was a homemaker, each with 50% probability. The randomized occupation of the father was either university professor (12.5% probability), an unskilled worker (12.5%), or an intermediate technical (37.5%) or commercial (37.5%) activity that matched the job type of the apprenticeship the application was sent to. Application gender was randomized with a 50% chance to be either female or male, independently of parental occupation. We considered 30 different job types for the application process, which we classified into female dominated (6), male dominated (16), and rather gender neutral categories (8). Furthermore, we classified job types classified w.r.t. three levels of requirements in terms of education and qualifications (low, average, and high), in order to investigate differences of employers' call back behaviour across job type classifications.

By and large, we find no robust evidence for discrimination based on applicant gender or parental occupation. For all but one of the investigated combinations of gender and occupational choice, differences in call back rates are not statistically significant at any conventional level when accounting for multiple hypothesis testing. The one exception is stating father's occupation to be university professor, which statistically significantly boots call backs for female applications, but not for male applications. Our results therefore provide some support for a blind recruitment procedure. Personal attributes (such parental occupation) should not be communicated to the employer in the first round of an application process, in order to prevent signalling effects and set the callback chances of all applicants on an equal footing.

Point estimates across subsamples suggest that the professor effect for female applications is to a larger extent driven by the German rather than the French speaking sample, by in terms of qualifications less rather than more demanding apprenticeships, and by more female- rather than male-dominated apprenticeships. However, due to low statistical power and issues related to multiple hypotheses testing we abstain from putting strong interpretations on the effect heterogeneities found across subsamples. The findings across subgroups generally back those of the main analysis. Specifically, when excluding the empirically rare case of having a professor as parent, we find no statistically significantly differential callback rates across gender and/or parental occupation that would suggest that employers contributed to occupational segregation.

Our paper is structured as follows. Section 2 reviews the literature on labor market discrimination and correspondence testing. Section 3 provides institutional background information on the Swiss educational system and apprenticeship market. Section 4 outlines the experimental design. Section 5 provides descriptive statistics for our data. Section 6 presents the empirical results. Section 7 concludes.

2 Literature Review on Labor Market Discrimination

Research on labor market discrimination has its foundation in three main studies which date back to the late 50's and early 70's. The seminal works of Becker (1957), who developed a model of taste for discrimination, as well as Arrow (1971) and Phelps (1972), who developed theories of statistical discrimination (regarding both sexism and racism), paved the way for studies on both ethnic and gender discrimination, many of them based on experimental methods.[1]

Riach and Rich (1995), for instance, find statistically significant discrimination against female applications for various job types in a correspondence test conducted in the Australian state of Victoria. For England, Riach and Rich (2006) document significantly lower call back rates for male applications in the case of secretary positions and for female applications in the case of engineering positions. We refer to Rich (2014), Baert (2017), and Neumark (2018) for thorough international reviews of experimental studies and correspondence tests investigating labor market discrimination and call back rates across genders in particular.

In the continental European context, Petit (2007) focuses on the effects of age and family on gender hiring decisions in the French financial sector. She finds significantly lower callback rates for younger women applying for high skilled administration jobs with long term contracts, but no discrimination among single and childless 37 year old applicants. For Sweden, Bygren et al. (2017) assess whether gender and parenthood status affects callback rates, and whether this is conditional on the qualifications required by the job applied for. They find no evidence for discrimination for neither the less nor more highly qualified jobs. Becker et al. (2018) conduct a correspondence test in Switzerland, Germany, and Austria to investigate how having a family affects callback rates for 32-year old applicants, but do not find statistically significant effects for females or males. However, women's chances of receiving a callback are reduced relative to men if living far from the workplace, applying to large companies, and having skills that do not match job requirements well.

While most studies focus on prime age workers, such that statistical discrimination related to family obligations could partly explain gender differences in callback

[1] Regarding ethnic discrimination, e.g. Bertrand and Mullainathan (2004), Jacquemet and Yannelis (2012), Carlsson and Rooth (2008), Carlsson and Rooth (2007), Veit and Yemane (2018) experimentally investigate how ethnicity revealed by (first or second) names.

rates, Kübler et al. (2018) similarly to our paper focus on the apprenticeship market. They run a vignette study in order to test for gender discrimination by sending short CVs to HR managers of German firms in order to have the applicants evaluated. The authors find that females are evaluated worse than males on average, but that the discrimination varies across industries and occupations. While parental occupation of the applicants is used as control variable, Kübler et al. (2018) do not assess the effect of parental occupation on callback rates as we do in this paper.

Furthermore, there exists a broad literature that focuses on parental education and/or occupation and its effect on health, education, occupations and/or further later life decisions of the offspring, see for instance Bello and Morchio (n.d.), Ham et al. (2009), Downey (1995), Giannelli and Rapallini (2018), Chevalier (2004)). However, to the best of our knowledge, there is no study which analyzes the impact of parental occupation on callback rates in the context of correspondence testing.

3 The Swiss Education System and Apprenticeship Market

In Switzerland, the responsibility for the educational system is mainly with the 26 cantons (regional administrative units), while the Swiss constitution only broadly defines the general foundations, like e.g. obligatorily free access to primary schooling. For this reason, there is considerable variation in school systems across cantons, albeit there are also attempts to harmonize key aspects of compulsory schooling through the so-called HarmoS concordate. The vast majority of students in Switzerland attend public schools close to their place of residence, only 5% go to private schools.

In most cantons (in particular those participating in the HarmoS concordate), compulsory schooling consists of 11 years of education, including two years of preschool or kindergarten attendance, starting at the age of four. After pre-school or kindergarten, primary schooling typically consists of 6 years, lower secondary schooling of 3 years. In the last year of primary school, students are assessed (also, but not exclusively based on their performance) in order to track them into three different types of lower secondary education that differ in terms of skill levels. After finishing lower secondary education, students enter, depending on the skill type accomplished, either the vocational education and training (VET) track, consisting of a dual apprenticeship system of formal education and training in a company, or the academic track, by attending either a general or specialized high school that prepare students for tertiary education.[2]

In Switzerland, roughly two thirds of all students with completed compulsory education enter the VET track and have around 230 occupations to choose from, see State Secretariat for Education Research and Innovation (2018)). An apprenticeship related to a particular occupation takes between three to four years. During this period, classes at a vocational school are combined with on-the-job training at a host company, where the apprentices are employed and paid a salary which increases with each completed year. It might also be the case, however, that students for a (fully) school-based VET program, which although less common overall, is particularly popular in the French and Italian speaking regions of Switzerland. Upon successful completion of the program, apprentices receive a federal VET diploma, which does not only serve as recognized occupational qualification, but also is a precondition for being eligible for receiving further education and higher qualifications in the chosen occupation. The VET system is managed as a public-private partnership, with the federal and cantonal governments as well as the companies and professional organisations jointly defining the curricula, skill sets and standards for occupations. Moreover, it is the companies that cover the costs for on the job training, salaries, and intracompany courses. The cantons, on the other hand, fund the vocational schools and career guidance services.

4 Experimental Design

Our correspondence test in the Swiss apprenticeship market consisted of a preparatory phase from October 2017 to July 2018 and an experimental phase from August 2018 until February 2019. In the preparatory phase, we first screened open apprenticeships as well as information about documents required in the application process[3] and consulted teenagers applying for apprenticeships to learn what typical applications look like. Specifically, we gathered typical CVs and motivation letters to use them as templates for our applications.

Furthermore, we classified apprenticeship types w.r.t. gender (non-)neutrality, relying on information about the relative popularity of specific occupations across genders provided online by the Educational Office of the Canton of Bern and the

[2]This section relies on the information provided by the State Secretariat for Education Research and Innovation (2018)

[3]Such information is, for instance, provided on the websites https://www.berufsberatung.ch and https://www.yousty.ch which we accessed in late 2017.

'Office for Equality of Males and Females' of the Canton of Zurich[4] and also cross-checked it with further online resources on the apprenticeship market.[5] Accordingly, we categorized occupations into clearly male-dominated, clearly female-dominated, and rather gender neutral types. We ultimately selected thirty occupations to be considered in the experiment, 8 of which are gender neutral (e.g. baker, cook, sales assistant, designer, etc), 6 female-dominated (e.g. hair dresser, dental assistant, medical practice assistant, etc.), and 16 male-dominated (e.g. gardener, carpenter, car mechanic, mason, electrician, etc.).

A second classification was w.r.t. the level of qualifications expected in terms of lower secondary schooling. We classified apprenticeship types into three levels of requirements, henceforth referred to as tiers, in order to adapt school certificates and aptitude tests accordingly to make applications look appropriate in terms of skills typically expected. For the first tier that is lowest in terms of requirements, applications contained lower secondary school certificates of the most basic skill type and comparably low scores from an aptitude test, if the latter was required at all, which was occupation-dependent. For the second tier, intermediate school certificates and aptitude test scores were used, while for the most demanding third tier, more advanced certificates and better test scores were included in the application documents.

Aiming to find an acceptable balance between expected sample sizes and organizational burden in preparing and managing applications, we ultimately decided to focus on 3 German speaking regions, namely the agglomerations of Basel, Bern, and Zurich and one French speaking region, the agglomeration of Lausanne. We prepared fictitious motivation letters, CVs, school certificates, and aptitude tests as well as two female and male profiles for either language region with varying names, addresses, and photos. Concerning names, we took the most popular choices for first names for either gender in 2004 in the German and French speaking parts, respectively, while the last names corresponded to the most frequent occurrences in the phone book in either language region. We also picked residential addresses

[4] See https://www.erz.be.ch/erz/de/index/berufsbildung/grundbildung/kennzahlen_berufsbildung/kennzahlen_berufsbildung2.html and https://ffg.zh.ch/internet/justiz_inneres/ffg/de/bildung/berufswahl/_jcr_content/contentPar/morethemes/morethemesitems/factsheet_die_belieb.spooler.download.1393238737874.pdf/FFG_2013_factsheet_die_beliebtesten_berufe_von_maedchen_und_jungen.pdf, respectively, both accessed in the beginning of 2018.

[5] See for instance the following list of the 10 most popular apprenticeships for females and males in 2015: https://blog.100000jobs.ch/de/2016/09/die-top-10-der-beliebtesten-lehrstellen/, accessed in the beginning of 2018.

in the 4 agglomerations for the fictitious candidates. Preparing school certificates that matched these addresses turned out to be more complicated than initially expected, first because certificates look different in each canton (and even over time) and second, because having to adapt certificates to skill levels appropriate for the 3 different tiers of apprenticeships. While applicant addresses and school certificates match in terms of cantonal congruence for Bern and Zurich, this is not the case for Basel and Lausanne. For the latter two agglomerations, it was apparent from the application documents that the respective fictitious candidate had recently moved from a different region.

Our aim was to send out two applications per open apprenticeship and to only consider one apprenticeship per employer in order to not overstrain companies with our experiment. In the CVs, applicant gender were randomized independently with a probability of 50% for each value. It was therefore possible that applications with both same or different genders were sent to a specific position, thus requiring 2 profiles per gender and language region. Also some further features like whether the applicant had a brother or sister were randomized this way. In contrast, mother's occupation was randomized pairwise among the two applications per position, implying that these applications had necessarily different vales for mother's occupation. The latter was either homemaker or primary school teacher, each with a chance of 50%.

Also father's occupation was randomized pairwise (and independently so of mother's occupation) and contained the following options: university professor (with 12.5% probability), an intermediate technical position (37.5%) matching the job type of the apprenticeship (e.g. mechanic), an intermediate commercial position (37.5%) matching the job type (e.g. sales manager), and an unskilled worker (12.5%). The idea was to consider high skilled, low skilled as well as intermediate profiles, with the latter being related to the position to be filled. The skill level of intermediate profiles therefore varied depending on the tier and industry of the position. For instance, for a technical apprenticeship in the first, second, or third tier, father's intermediate technical occupation would either be a mechanic, a polymechanic, or an engineer. This implies substantial heterogeneity of educational achievements within the intermediate profiles for the sake of aligning father's occupation well with open apprenticeships. Some further CV features like motivational sentences and leisure activities were also randomized pairwise in order to make sure that not the same phrases are used in two applications sent to the same vacancy.

In total, 3069 applications were sent out between August and mid October 2018 via e-mail to open positions posted on Switzerland's most popular online portal for apprenticeships. During the process, several issues arose. In August, we accidentally sent out applications also to some positions that were from the previous year and thus not relevant for our fictitious candidates. In a few cases, the employers' e-mail addresses provided online contained typos or were not actual such that the applications could not be sent. 129 observations were dropped due to such issues. A more serious concern was that 5 employers in the German speaking part detected that our applications were not related to existing students, by following up on the candidates by consulting the schools. Even though these cases were excluded from the analysis, it cannot be ruled out that the information was spread to further employers. This would in the worst case bias any effects towards zero by ignoring any of our applications. However, robustness checks presented further below do not suggest that these issues affected overall callback rates. Furthermore, in 1 case (German speaking part), we unintendedly sent out 4 applications to the same employer such that the same applicant name occurred twice. Even though we did not get any negative reaction, we immediately withdrew our applications when noticing the issue and excluded this employer, too. All in all, we dropped 12 observations because issues mentioned. Our final evaluation data set thus consists of 2928 observations. While most employers received two applications as intended, 397 employers in Lausanne only received 1 application, due to organizational issues at the end of the application period (end of September until mid October).

Employers mainly responded via e-mails, but frequently also via phone calls to the voicemail boxes that we had set up for each candidate profile. Much more rarely, we received physical letters at the addresses indicated in the fictitious CVs. In 10 instances, such letters could not be delivered and were returned to employers, who then wrote e-mails to ask for a correct address. In these cases, we replaced the problematic addresses (also for any further applications) and apologized via e-mail and also asked to send the letters to the new address or answer via e-mail instead. These employers are kept in our evaluation sample, albeit excluding them leaves our results virtually unchanged. In the case that one of our applications received an invitation, which was either to a job interview, assessment center, or trial apprenticeship, we declined the offer within several days. In this case, the dependent variable, employer response, was coded as one. In the case of a negative response or no reaction of the employer up to February 2019, the dependent variable was coded as zero.

5 Data and Descriptive Statistics

Our evaluation sample consists of 2928 observations and contains a range of application-related characteristics, like the previously mentioned apprenticeship tiers and classification in terms of gender neutrality, applicant gender and city, parental occupation, and a dummy for whether the applicant had moved with school certificates from a different region. Furthermore, we relative to August 1st 2018 recorded the dates when the apprenticeships were posted or (in case this was unclear) found by us, as well as the dates when we sent out the applications. During the application process, we also collected information on employers, namely (an estimation of the) number of employees, sector (public, trade and wholesale, manufacturing and goods, or services), geographic distance to the central station of the applicant's city, scale of activity (local, national, or international), explicit indication of an anti-discrimination policy on the website, and the gender of the contact person in the company(female, male, or unknown).

Table 1 provides the means of all characteristics but gender in the total sample as well as separately for females and males. It also contains mean differences across gender ('diff') and p-values ('p-val') of two sample t-tests. The means of the variables are generally well balanced across gender, and mean differences that are statistically significant at the 5% level are the exception. For testing mean balance of all characteristics jointly, we the apply machine learning-based test suggested by Ludwig et al. (2017). It is based on the intuition that the problem of obtaining too many significant results when testing multiple hypotheses (e.g. mean differences in multiple characteristics across gender), or false positives, is similar to the concern of overfitting in machine learning.

We thus follow Ludwig et al. (2017) and apply the machine learning logic by splitting our sample into training and testing data. In the training data, we run a lasso logit regression of gender on the characteristics using the 'rlogit' command with its default values in the 'hdm' package by Chernozhukov et al. (2015) for the statistical software 'R'. We then use the obtained coefficients for predicting gender in the test data and compare the prediction to actual gender to compute the mean squared error (MSE). We use 5-fold cross-validation, such that the roles of training and test data are swapped, and take the average of the 5 MSEs obtained. In a

Table 1: Descriptive statistics by applicant gender

	total sample	female	male	t-test	
	mean	mean	mean	diff	p-val
employees: 1 to 20	0.48	0.48	0.47	0.00	0.88
employees: 21 to 50	0.24	0.24	0.24	-0.01	0.63
employees: 51 to 100	0.11	0.11	0.12	-0.00	0.74
employees: 101 to 250	0.09	0.09	0.08	0.01	0.52
employes: 251 to 500	0.03	0.03	0.03	-0.00	0.87
employees: 501 to 1000	0.02	0.02	0.02	-0.00	0.97
employees: more than 1000	0.04	0.04	0.04	0.00	0.67
sector: public	0.05	0.05	0.06	-0.00	0.69
sector: trade and wholesale	0.22	0.22	0.23	-0.01	0.48
sector: manufacturing and goods	0.13	0.12	0.13	-0.01	0.64
sector: services	0.60	0.61	0.59	0.02	0.27
distance to city center	16.37	16.22	16.55	-0.33	0.57
tier 1 job	0.35	0.34	0.36	-0.02	0.32
tier 2 job	0.36	0.38	0.35	0.03	0.15
tier 3 job	0.28	0.28	0.29	-0.01	0.63
type: gender-neutral	0.33	0.32	0.33	-0.01	0.56
type: female-dominated	0.25	0.27	0.22	0.05	0.00
type: male-dominated	0.43	0.41	0.45	-0.04	0.03
city: Bern	0.21	0.21	0.21	-0.00	0.91
city: Zurich	0.30	0.29	0.31	-0.02	0.30
city: Basel	0.11	0.13	0.09	0.03	0.00
city: Lausanne	0.38	0.37	0.39	-0.02	0.39
activity: regional	0.80	0.81	0.79	0.02	0.27
activity: national	0.12	0.11	0.13	-0.02	0.16
activity: international	0.08	0.08	0.08	0.00	0.95
antidiscrimination policy	0.21	0.22	0.20	0.02	0.23
contact: female	0.31	0.32	0.30	0.02	0.29
contact: male	0.33	0.32	0.34	-0.02	0.36
contact: unknown	0.36	0.36	0.36	-0.00	0.90
day job was published or found	29.08	29.00	29.17	-0.18	0.74
day of application	51.00	50.80	51.22	-0.42	0.50
father professor	0.13	0.12	0.14	-0.03	0.04
father intermediate	0.75	0.76	0.74	0.02	0.15
father unskilled worker	0.12	0.12	0.12	0.00	0.82
mother teacher	0.50	0.51	0.48	0.03	0.07
applicant has moved	0.49	0.50	0.48	0.02	0.29
number of observations	2928	1529	1399		

Note: Means of characteristics in the total, female, and male samples, as well as mean differences ('diff') between females and males and p-values ('p-val')

next step, we randomly relabel (or permute)gender and re-estimate the MSE using the same procedure. Repeating the permutation 999 times, we compute the p-value for the joint significance of the characteristics as the share of permutation based MSEs that are lower than the MSE with the correct coding of the treatment. The permutation test's intuition is that if the characteristics are balanced across gender, relabelling will not seriously affect (i.e. increase) the MSE. If, on the other hand, characteristics are predictive for gender, the correct coding of gender should likely entail a smaller MSE than the permuted versions. The p-value of the test is 0.984 (or 98.4%), thus providing no evidence for joint imbalances.

Table 2: Descriptive statistics by parental occupation

	m_te_f_un	m_te_f_in		m_te_f_pr		m_ho_f_un		m_ho_f_in		m_ho_f_pr	
	mean	diff	p-val	diff	p-val	diff	p-val	diff	p-val	diff	p-val
employees: 1 to 20	0.56	-0.09	0.03	-0.17	0.00	-0.10	0.07	-0.08	0.06	-0.13	0.02
employees: 21 to 50	0.22	0.03	0.46	0.01	0.89	0.02	0.61	0.01	0.74	0.03	0.47
employees: 51 to 100	0.09	0.02	0.30	0.07	0.04	0.05	0.16	0.03	0.20	-0.01	0.74
employees: 101 to 250	0.06	0.03	0.13	0.05	0.08	0.00	0.98	0.03	0.13	0.08	0.01
employes: 251 to 500	0.04	-0.02	0.16	0.00	0.91	0.00	0.90	-0.02	0.28	-0.03	0.13
employees: 501 to 1000	0.01	0.01	0.14	0.03	0.06	0.01	0.23	0.01	0.27	0.02	0.13
employees: more than 1000	0.02	0.02	0.12	0.00	0.83	0.01	0.73	0.02	0.21	0.03	0.13
sector: public	0.02	0.04	0.00	0.03	0.10	0.07	0.00	0.03	0.01	0.04	0.03
sector: trade and wholesale	0.21	0.01	0.78	-0.01	0.83	0.02	0.65	0.03	0.44	-0.03	0.54
sector: manufacturing and goods	0.16	-0.03	0.29	-0.06	0.12	-0.05	0.19	-0.03	0.28	-0.02	0.65
sector: services	0.61	-0.01	0.74	0.03	0.52	-0.04	0.44	-0.02	0.56	0.00	0.99
distance to city center	18.96	-2.72	0.04	-3.29	0.05	-2.28	0.18	-3.03	0.02	-1.22	0.46
tier 1 job	0.36	-0.02	0.61	0.05	0.38	-0.01	0.82	-0.01	0.76	0.02	0.64
tier 2 job	0.37	-0.01	0.88	-0.03	0.60	0.00	0.98	-0.01	0.77	-0.04	0.38
tier 3 job	0.26	0.03	0.48	-0.02	0.68	0.01	0.83	0.02	0.52	0.02	0.67
type: gender-neutral	0.30	0.02	0.63	0.05	0.31	0.01	0.77	0.04	0.29	0.03	0.49
type: female-dominated	0.25	-0.01	0.84	-0.00	0.97	0.00	0.96	-0.01	0.83	-0.01	0.86
type: male-dominated	0.45	-0.01	0.80	-0.05	0.35	-0.02	0.76	-0.03	0.43	-0.03	0.62
city: Bern	0.21	0.00	0.90	0.00	0.97	-0.00	0.99	-0.01	0.88	0.04	0.43
city: Zurich	0.31	-0.02	0.66	0.01	0.92	-0.02	0.71	-0.01	0.73	-0.01	0.79
city: Basel	0.10	0.00	0.99	0.02	0.55	0.03	0.42	0.01	0.84	0.00	0.94
city: Lausanne	0.37	0.01	0.76	-0.03	0.60	-0.01	0.86	0.01	0.75	-0.02	0.64
activity: regional	0.83	-0.03	0.36	-0.04	0.36	-0.03	0.50	-0.04	0.26	-0.05	0.25
activity: national	0.11	0.00	0.88	0.03	0.38	0.01	0.85	0.01	0.62	0.01	0.74
activity: international	0.06	0.03	0.20	0.01	0.78	0.02	0.42	0.02	0.25	0.04	0.19
antidiscrimination policy	0.17	0.04	0.18	0.06	0.16	0.01	0.88	0.04	0.22	0.07	0.09
contact: female	0.26	0.05	0.16	0.05	0.27	0.04	0.46	0.05	0.19	0.05	0.29
contact: male	0.34	-0.01	0.76	-0.02	0.70	0.02	0.74	-0.02	0.65	-0.00	0.95
contact: unknown	0.39	-0.04	0.32	-0.03	0.51	-0.05	0.31	-0.03	0.45	-0.05	0.36
day job was published or found	28.44	0.66	0.60	0.82	0.62	0.18	0.91	1.04	0.41	-0.79	0.61
day of application	49.45	1.98	0.17	-0.48	0.80	0.80	0.66	2.38	0.10	-1.66	0.37
applicant: female	0.55	0.00	0.98	-0.06	0.25	-0.03	0.53	-0.03	0.42	-0.08	0.11
applicant has moved	0.48	0.01	0.76	-0.01	0.90	0.02	0.72	0.02	0.66	-0.02	0.68
number of observations	163	1119		176		197		1076		197	

Note: 'm_te_f_un' provides the means of characteristics in the reference group (mother teacher, father unskilled worker), the other columns provide the mean differences ('diff') compared to the baseline group and the p-values ('p-val'), respectively. 'm_te_f_in': mother teacher, father intermediate; 'm_te_f_pr': mother teacher, father professor; 'm_ho_f_un': mother homemaker, father unskilled worker; 'm_ho_f_in': mother homemaker, father intermediate; 'm_ho_f_pr': mother homemaker, father professor.

Table 2 reports the means of all characteristics but parental occupation for the group of applications with the mother being a teacher and the father being an unskilled worker ('mean'). Furthermore, it shows mean differences ('diff') between this reference group and other combinations of parental occupation, namely: mother is a teacher and father has an intermediate occupation (technical or commercial), mother is teacher and father is a university professor, mother is a homemaker and father is a low skilled worker, mother is a homemaker and father has an intermediate occupation (technical or commercial), and mother is homemaker and father is a university professor. P-values for the respective two sample t-tests are also reported ('p-val'). Again, the majority of mean differences is not statistically significant at the 5%. We also apply the procedure of Ludwig et al. (2017) for pairwise testing of mother is a teacher vs. mother is a homemaker, father has an intermediate occupation vs. father has a different occupation, and father is a professor vs. father is not a professor. The respective p-values are 5.2%, 91.6%, and 96.7%. By and large, characteristics thus appear satisfactorily balanced across our intervention variables of interest, namely applicant gender and parental occupation. For the single variable of mother's occupation, however, balance is almost rejected at the 5% level of significance, but this p-value does not account for the fact that we run the Ludwig et al. (2017) test for multiple hypotheses. In any case, our empirical results presented in Section 6 are very similar when conditioning or not conditioning for application and employer characteristics to control for observed imbalances.

6 Results

In our main analysis, we run a fully saturated linear regression of the dependent variable employer response (1 for invitation and 0 for no invitation) on dummies for each possible combination of applicant gender (female or male) and parental occupation (mother teacher and father worker, mother teacher and father intermediate, mother teacher and father professor, mother homemaker and father worker, mother homemaker and father intermediate, mother homemaker and father professor). Standard errors are computed by cluster-bootstrapping the coefficients, where clustering is on the employer-level.

Table 3 reports the results. The reference category are female applicants with mothers working as teachers and fathers being unskilled workers, for which the average employer response (i.e. the share of invitations) is reported ('est'), which amounts to roughly 19%. For the other 11 categories defined by combinations of gender and parental occupation, we report the respective difference to the reference category ('est'), along with bootstrap standard errors ('boot se') and conventional p-values ('raw p-val') based on a t-test. However, these p-values do not take into account for multiple hypothesis testing, i.e. the fact that we simultaneously test 11 differences. This is problematic because the likelihood of spuriously rejecting one or even several null hypotheses generally increases in the number of hypotheses tested. We therefore apply the approach of Romano and Wolf (2016) to adjust the p-values of each difference for multiple testing ('adj p-val').

We find that when accounting for multiple testing, most differences relative to the reference category are statistically insignificant at conventional levels. One exception is having a professor as father, which boosts callback rates for females by more than

Table 3: Effects of gender and parental occupation

	est	boot se	raw p-val	adj p-val
female: mother teacher, father unskilled worker (mean)	0.191	0.042	0.000	
female: mother teacher, father intermediate	0.131	0.047	0.005	0.076
female: mother teacher, father professor	0.221	0.070	0.002	0.004
female: mother home, father unskilled worker	0.096	0.062	0.118	0.216
female: mother home, father intermediate	0.088	0.044	0.045	0.216
female: mother home, father professor	0.205	0.064	0.001	0.004
male: mother teacher, father unskilled worker	0.066	0.066	0.318	0.280
male: mother teacher, father intermediate	0.093	0.046	0.043	0.216
male: mother teacher, father professor	0.062	0.062	0.320	0.280
male: mother home, father unskilled worker	0.090	0.063	0.151	0.216
male: mother home, father intermediate	0.074	0.045	0.103	0.280
male: mother home, father professor	0.073	0.060	0.226	0.280
number of observations	2928			

Note: estimates of (differences in) callback rates for the total sample, without control variables. 'est' provides the callback rate for the group 'female: mother teacher, father intermediate', as well as the differences in callback rates of all other groups relative to 'female: mother teacher, father intermediate'. 'boot se' reports bootstrap standard errors clustered at the employer level. 'raw p-val' gives the p-values not accounting for multiple hypothesis testing. 'adj p-val' provides adjusted p-values accounting for multiple hypothesis testing

20 percentage points, independent of mother's occupation. The effect of either category with female applicants and professors as fathers is statistically significant at the 1% level, even when accounting for multiple testing. Furthermore, having a mother who is a teacher and a father with an intermediate occupation also increases callback rates for females relative to the reference group and is significant at the 10% level. In contrast, callback rates of male applications are rather stable and not significantly different across parental occupation.

As a robustness check, we include the applicant and employer characteristics reported in Tables 1 and 2 as control variables to account for observed imbalances. This does not importantly change our findings, see Table 4. The effect of having a professor as father among female applications remains large (roughly 20 percentage points) and statistically significant at the 5%. For any other combination of applicant gender and parental occupation, dofferences to the reference group are not statistically significant at the 10% level. With the exception of the interaction between a female application and having a professor as father, we therefore find no robust statistical evidence for a systematically differential treatment based on gender or parental occupation. This concerns in particular empirically more relevant parental occupations that exclude the rare case of a professor. However, the estimates also suggest that parental occupation might have a signalling effect for female

Table 4: Treatment effects of gender and parental occupation with controls

	est	boot se	raw pval	adj pval
female: mother teacher, father unskilled worker (intercept)	0.323	0.087	0.000	
female: mother teacher, father intermediate	0.131	0.045	0.004	0.142
female: mother teacher, father professor	0.211	0.064	0.001	0.007
female: mother home, father unskilled worker	0.090	0.058	0.122	0.406
female: mother home, father intermediate	0.084	0.043	0.050	0.455
female: mother home, father professor	0.189	0.061	0.002	0.013
male: mother teacher, father unskilled worker	0.082	0.065	0.209	0.471
male: mother teacher, father intermediate	0.079	0.045	0.080	0.473
male: mother teacher, father professor	0.052	0.059	0.382	0.763
male: mother home, father unskilled worker	0.092	0.061	0.131	0.404
male: mother home, father intermediate	0.069	0.044	0.115	0.571
male: mother home, father professor	0.050	0.058	0.388	0.763
number of observations	2928			

Note: estimates of (differences in) callback rates for the total sample, with control variables. 'est' provides the callback rate for the group 'female: mother teacher, father intermediate', as well as the differences in callback rates of all other groups relative to 'female: mother teacher, father intermediate'. 'boot se' reports bootstrap standard errors clustered at the employer level. 'raw p-val' gives the p-values not accounting for multiple hypothesis testing. 'adj p-val' provides adjusted p-values accounting for multiple hypothesis testing.

applications, which is absent among males.

Table 5 provides a more parsimonious way to display our main findings. It presents differences in callback rates between females and males by whether the father is (not) a professor. In the subsample without professor, females are 2 percentage points more likely to be invited than males (see the upper left panel), but the difference is far from being statistically significant. Results are very similar when controlling for applicant and employer characteristics (upper right panel). In contrast, in the subsample with a professor as father (lower panels), females are more than 12 percentage points more likely to be invited than males. This difference is statistically significant at the 5% level even when accounting for multiple hypothesis testing of the coefficients on gender, professor, and the gender-professor-interaction.

Next, we investigate the heterogeneity of our results across language regions. For this reason, we separately run the analysis for the German (Basel, Bern, and Zurich) and the French (Lausanne) speaking regions to explore relative effect sizes, see Table 6 in the Appendix. The professor effects among female applicants are positive in either language group, but on average larger in the German speaking sample. However, it cannot be rejected at conventional levels of significance that the respective estimates in the French and German speaking samples are actually the same. While the professor effect for females is to a larger extent driven by the

Table 5: Gender differences in callbacks by professor status

	est	boot se	raw pval	adj pval	est	boot se	raw pval	adj pval
father is not professor	no controls				with controls			
male (mean / intercept)	0.274				0.397			
female (diff)	0.020	0.019	0.301	0.348	0.024	0.017	0.157	0.268
number of observations	2555				2555			
father is professor	no controls				with controls			
male (mean / intercept)	0.259				0.373			
female (diff)	0.125	0.053	0.018	0.013	0.125	0.050	0.013	0.006
number of observations	373				373			

Note: estimates of (differences in) callback rates across gender when father is reported or not reported to be a professor, without and with control variables. 'est' provides the callback rates for males and the difference in call back rates between females and males. 'boot se' reports bootstrap standard errors clustered at the employer level. 'raw p-val' gives the p-values not accounting for multiple hypothesis testing. 'adj p-val' provides adjusted p-values accounting for multiple hypothesis testing of differences by gender, professor, and the gender-professor-interaction.

German speaking region in terms of point estimates, statistical power is too low to make clear cut conclusion across language regions, in particular when accounting for multiple hypothesis testing issues introduced by splitting by language region. Any other difference in callback rates relative to the reference group is insignificant in either language group.

We conduct two further heterogeneity checks by either running the analysis within tiers (i.e. levels of qualification), see Table 7 in the Appendix, or within female-dominated, gender neutral, and male-dominated apprenticeship types, see Table 8 in the Appendix. It appears that the lower tiers 1 and 2 as well as the female-dominated and gender neutral apprenticeships drive the female-professor interaction effect found in the main sample. However, we abstain from making strong claims about differences across subgroups, due to low statistical power and issues of multiple hypothesis testing. By and large, the point estimates support our main findings in the total sample.

In Section 4 we discussed that to the best of our knowledge 5 employers detected that our applications were not related to existing students. 4 detections were related to applications sent out between August 28th and September 7th, only 1 detection to applications in October. As a robustness check, we therefore run our analysis for the month September only, to investigate whether a potential communication among employers about the detection of fictitious applications affected our main findings. Even though we cannot rule out that some employers exchanged information on this issue and adapted their response behaviour accordingly, our results do not suggest that this is a widespread phenomenon. As can be seen in Table 9 in the Appendix, the results are qualitatively in line with those based on the total sample. In the case of a thorough dissemination of information on our experiment, in particular the applicant names, one would expect the share of positive responses to be close to zero for any of our profiles. However, callback rates are generally far from zero and statistically significant and also the female-professor-interaction effect is quantitatively not too different to that in the main sample, albeit estimated less precisely.

7 Conclusion

We investigated the effects of gender and parental occupation on callback rates for applications to apprenticeships by means of a correspondence test in Switzerland. Sending out roughly 3000 fictitious applications in 4 different regions, our intervention variables did not statistically significantly affect callbacks in most cases. We therefore found no robust evidence for a differential treatment of applications w.r.t. to gender or parental occupation in the Swiss apprenticeship market. The one exception was stating university professor as father's occupation, which statistically significantly boosted callbacks for females even when accounting for multiple hypothesis testing, but not for males. This suggests that applications should ideally be blind and not reveal socio-economic information to maximize fairness. Our findings remained qualitatively unchanged when controlling for observed characteristics of applications and employers or when considering subsamples defined upon language regions, apprenticeship types, or the timing of application.

References

Arrow, Kenneth, "The Theory of Discrimination," Working Papers 403, Princeton University, Department of Economics, Industrial Relations Section. October 1971.

Baert, Stijn, "Hiring Discrimination: An Overview of (Almost) all Correspondence Experiments Since 2005," *GLO Discussion Paper 61*, 2017.

Becker, Gary Stanley, *The Economics of discrimination*, University of Chicago Press, 1957.

Becker, Sascha O, Ana Fernandes, and Doris Weichselbaumer, "Discrimination Based on Potential and Realized Fertility in Hiring: Evidence from a Large-Scale Field Experiment," *working paper, Bern University of Applied Sciences*, 2018.

Bello, Salvatore Lo and Iacopo Morchio, "Parental Links and Labor Market Outcomes : Evidence from the UK."

Bertrand, Marianne and Sendhil Mullainathan, "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination," *American Economic Review*, 2004, *94* (4), 991–1013.

Blau, Francine D. and Lawrence M. Kahn, "The Gender Wage Gap: Extent, Trends, and Explanations," *Journal of Economic Literature*, 2017, *55* (3), 789–865.

Bygren, Magnus, Anni Erlandsson, and Michael Gähler, "Do Employers Prefer Fathers? Evidence from a Field Experiment Testing the Gender by Parenthood Interaction Effect on Callbacks to Job Applications," *European Sociological Review*, 2017, *33* (3), 337–348.

Carlsson, Magnus and Dan Olof Rooth, "Evidence of ethnic discrimination in the Swedish labor market using experimental data," *Labour Economics*, 2007, *14* (4 SPEC. ISS.), 716–729.

_ **and Dan-Olof Rooth**, "Is it Your Foreign Name or Foreign Qualifications? An Experimental Study of Ethnic Discrimination in Hiring," *IZA Discussion Paper 3810*, 2008.

Chernozhukov, V., C. Hansen, and M. Spindler, "hdm: High-dimensional metrics," *A package for the statistical software R*, 2015.

Chevalier, Arnaud, "Parental Education and Child's Education: A Natural Experiment," *IZA Discussion Paper 1153*, 2004.

Cortes, Patricia and Jessica Pan, "Occupation and gender," in Susan Averett, Laura M. Argys, and Saul D. Hoffman, eds., *Susan Averett, Laura M. Argys, and Saul D. Hoffman, eds.*, Oxford University Press Oxford, UK 2018, p. 425–452.

Downey, Douglas B, "When Bigger Is Not Better: Family Size, Parental Resources, and Children's Educational Performance," *American Sociological Review*, 1995, *60* (5), 746–761.

Giannelli, Gianna C and Chiara Rapallini, "Parental Occupation and Children's School Outcomes in Math," *IZA Discussion Paper 11395*, 2018.

Ham, Roger, Pnr Junankar, and Robert Wells, "Occupational Choice: Personality Matters," *IZA Discussion Paper 4105*, 2009.

Huggett, Mark, Gustavo Ventura, and Amir Yaron, "Sources of Lifetime Inequality," *American Economic Review*, 2011, *101* (7), 2923–2954.

Jacquemet, Nicolas and Constantine Yannelis, "Indiscriminate discrimination: A correspondence test for ethnic homophily in the Chicago labor market," *Labour Economics*, 2012, *19* (6), 824–832.

Kübler, Dorothea, Julia Schmid, and Robert Stüber, "Gender discrimination in hiring across occupations: a nationally-representative vignette study," *Labour Economics*, 2018, *55*, 215–229.

Ludwig, J., S. Mullainathan, and J. Spiess, "Machine learning tests for effects on multiple outcomes," *working paper, Harvard University*, 2017.

Neumark, David, "Experimental Research on Labor Market Discrimination," *Journal of Economic Literature*, September 2018, *56* (3), 799–866.

Petit, Pascale, "The effects of age and family constraints on gender hiring discrimination: A field experiment in the French financial sector," *Labour Economics*, 2007, *14* (3), 371–391.

Phelps, Edmund S., "The Statistical Theory of Racism and Sexism," *The American Economic Review*, 1972, *62* (4), 659–661.

Riach, Peter A. and Judith Rich, "An Investigation of Gender Discrimination in Labor Hiring," *Eastern Economic Journal*, 1995, *21* (3), 343–356.

Riach, Peter A and Judith Rich, "An Experimental Investigation of Sexual Discrimination in Hiring in the English Labor Market," *Advances in Economic Analysis & Policy*, 2006, *5* (2), 1–20.

Rich, Judith, "What Do Field Experiments of Discrimination in Markets Tell Us? A Meta Analysis of Studies Conducted Since 2000," *IZA Discussion Paper 8584*, 2014, pp. 1–71.

Romano, Joseph P. and Michael Wolf, "Efficient computation of adjusted p-values for resampling-based stepdown multiple testing," *Statistics and Probability Letters*, 2016, *113*, 38–40.

State Secretariat for Education Research and Innovation, "Vocational and Professional Education and Training in Switzerland: Facts and Figures 2018," Technical Report, State Secretariat for Education, Research and Innovation 2018.

Veit, Susanne and Ruta Yemane, "The ADIS study: A large-scale correspondence test on labor market discrimination in Germany - Technical Report," Technical Report No. SP VI 2018-103 2018.

Weichselbaumer, Doris, "Is It Sex or Personality? The Impact of Sex Stereotypes on Discrimination in Applicant Selection," *Eastern Economic Journal*, 2004, *30* (2), 159–186.

A Appendix

Table 6: Treatment effects of gender and parental occupation across language regions

	est	boot se	raw pval	adj pval
German language region				
female: mother teacher, father unskilled worker (mean)	0.232	0.056	0.000	
female: mother teacher, father intermediate	0.112	0.062	0.069	0.241
female: mother teacher, father professor	0.222	0.088	0.011	0.025
female: mother home, father unskilled worker	0.127	0.084	0.129	0.206
female: mother home, father intermediate	0.096	0.058	0.098	0.323
female: mother home, father professor	0.259	0.086	0.003	0.008
male: mother teacher, father unskilled worker	0.072	0.089	0.416	0.461
male: mother teacher, father intermediate	0.118	0.062	0.059	0.215
male: mother teacher, father professor	0.068	0.082	0.407	0.461
male: mother home, father unskilled worker	0.030	0.078	0.699	0.484
male: mother home, father intermediate	0.064	0.058	0.271	0.461
male: mother home, father professor	0.035	0.078	0.651	0.484
number of observations	1815			
French language region				
female: mother teacher, father unskilled worker (mean)	0.121	0.059	0.038	
female: mother teacher, father intermediate	0.161	0.066	0.014	0.170
female: mother teacher, father professor	0.212	0.107	0.047	0.075
female: mother home, father unskilled worker	0.041	0.083	0.624	0.496
female: mother home, father intermediate	0.081	0.064	0.206	0.419
female: mother home, father professor	0.114	0.093	0.222	0.309
male: mother teacher, father unskilled worker	0.057	0.094	0.541	0.496
male: mother teacher, father intermediate	0.069	0.065	0.284	0.453
male: mother teacher, father professor	0.040	0.089	0.652	0.496
male: mother home, father unskilled worker	0.193	0.098	0.049	0.097
male: mother home, father intermediate	0.094	0.064	0.142	0.349
male: mother home, father professor	0.136	0.096	0.159	0.253
number of observations	1113			

Note: estimates of (differences in) callback rates per language region, without control variables. 'est' provides the callback rate for the group 'female: mother teacher, father intermediate', as well as the differences in callback rates of all other groups relative to 'female: mother teacher, father intermediate'. 'boot se' reports bootstrap standard errors clustered at the employer level. 'raw p-val' gives the p-values not accounting for multiple hypothesis testing. 'adj p-val' provides adjusted p-values accounting for multiple hypothesis testing.

Table 7: Treatment effects of gender and parental occupation across tiers

	est	boot se	raw pval	adj pval
Tiers 1 and 2				
female: mother teacher, father unskilled worker (mean)	0.131	0.042	0.002	
female: mother teacher, father intermediate	0.152	0.046	0.001	0.059
female: mother teacher, father professor	0.278	0.073	0.000	0.001
female: mother home, father unskilled worker	0.108	0.068	0.111	0.131
female: mother home, father intermediate	0.108	0.045	0.017	0.131
female: mother home, father professor	0.275	0.073	0.000	0.001
male: mother teacher, father unskilled worker	0.089	0.069	0.198	0.147
male: mother teacher, father intermediate	0.116	0.047	0.014	0.111
male: mother teacher, father professor	0.123	0.066	0.063	0.111
male: mother home, father unskilled worker	0.133	0.066	0.043	0.106
male: mother home, father intermediate	0.086	0.045	0.055	0.147
male: mother home, father professor	0.090	0.062	0.151	0.147
number of observations	2097			
Tier 3				
female: mother teacher, father unskilled worker (mean)	0.321	0.086	0.000	
female: mother teacher, father intermediate	0.102	0.095	0.284	0.626
female: mother teacher, father professor	0.100	0.151	0.510	0.641
female: mother home, father unskilled worker	0.079	0.125	0.531	0.676
female: mother home, father intermediate	0.062	0.090	0.492	0.676
female: mother home, father professor	0.049	0.120	0.683	0.676
male: mother teacher, father unskilled worker	0.079	0.154	0.610	0.676
male: mother teacher, father intermediate	0.046	0.095	0.626	0.676
male: mother teacher, father professor	-0.071	0.123	0.563	0.713
male: mother home, father unskilled worker	0.012	0.129	0.927	0.676
male: mother home, father intermediate	0.057	0.095	0.550	0.676
male: mother home, father professor	0.058	0.128	0.651	0.676
number of observations	831			

Note: estimates of (differences in) callback rates for tier 1 and 2 vs. tier 3, without control variables. 'est' provides the callback rate for the group 'female: mother teacher, father intermediate', as well as the differences in callback rates of all other groups relative to 'female: mother teacher, father intermediate'. 'boot se' reports bootstrap standard errors clustered at the employer level. 'raw p-val' gives the p-values not accounting for multiple hypothesis testing. 'adj p-val' provides adjusted p-values accounting for multiple hypothesis testing.

Table 8: Treatment effects of gender and parental occupation across types

	est	boot se	raw pval	adj pval
Female dominated apprenticeship types				
female: mother teacher, father unskilled worker (mean)	0.120	0.064	0.063	
female: mother teacher, father intermediate	0.137	0.072	0.057	0.167
female: mother teacher, father professor	0.315	0.121	0.009	0.022
female: mother home, father unskilled worker	0.065	0.101	0.520	0.469
female: mother home, father intermediate	0.093	0.075	0.217	0.363
female: mother home, father professor	0.213	0.107	0.047	0.122
male: mother teacher, father unskilled worker	0.192	0.136	0.156	0.144
male: mother teacher, father intermediate	0.027	0.073	0.710	0.574
male: mother teacher, father professor	-0.072	0.080	0.364	0.808
male: mother home, father unskilled worker	0.184	0.115	0.110	0.144
male: mother home, father intermediate	0.070	0.070	0.317	0.460
male: mother home, father professor	0.005	0.092	0.957	0.631
number of observations	718			
Gender neutral apprenticeship types				
female: mother teacher, father unskilled worker (mean)	0.182	0.082	0.026	
female: mother teacher, father intermediate	0.179	0.090	0.046	0.200
female: mother teacher, father professor	0.232	0.124	0.061	0.120
female: mother home, father unskilled worker	0.161	0.114	0.159	0.262
female: mother home, father intermediate	0.093	0.086	0.277	0.451
female: mother home, father professor	0.218	0.118	0.064	0.128
male: mother teacher, father unskilled worker	0.040	0.115	0.726	0.469
male: mother teacher, father intermediate	0.131	0.089	0.141	0.325
male: mother teacher, father professor	0.061	0.112	0.589	0.469
male: mother home, father unskilled worker	0.077	0.117	0.509	0.469
male: mother home, father intermediate	0.086	0.089	0.331	0.451
male: mother home, father professor	0.096	0.111	0.389	0.451
number of observations	964			
Male dominated apprenticeship types				
female: mother teacher, father unskilled worker (mean)	0.238	0.067	0.000	
female: mother teacher, father intermediate	0.098	0.075	0.189	0.433
female: mother teacher, father professor	0.156	0.110	0.156	0.259
female: mother home, father unskilled worker	0.070	0.104	0.502	0.506
female: mother home, father intermediate	0.088	0.069	0.201	0.476
female: mother home, father professor	0.194	0.098	0.048	0.156
male: mother teacher, father unskilled worker	0.020	0.100	0.842	0.506
male: mother teacher, father intermediate	0.084	0.075	0.259	0.476
male: mother teacher, father professor	0.140	0.102	0.170	0.273
male: mother home, father unskilled worker	0.045	0.098	0.648	0.506
male: mother home, father intermediate	0.064	0.073	0.379	0.506
male: mother home, father professor	0.088	0.095	0.355	0.476
number of observations	1246			

Note: estimates of (differences in) callback rates for tier 1 and 2 vs. tier 3, without control variables. 'est' provides the callback rate for the group 'female: mother teacher, father intermediate', as well as the differences in callback rates of all other groups relative to 'female: mother teacher, father intermediate'. 'boot se' reports bootstrap standard errors clustered at the employer level. 'raw p-val' gives the p-values not accounting for multiple hypothesis testing. 'adj p-val' provides adjusted p-values accounting for multiple hypothesis testing.

Table 9: Treatment effects of gender and parental occupation in September 2018

	est	boot se	raw pval	adj pval
female: mother teacher, father unskilled worker (mean)	0.184	0.063	0.004	
female: mother teacher, father intermediate	0.118	0.069	0.088	0.272
female: mother teacher, father professor	0.248	0.102	0.014	0.035
female: mother home, father unskilled worker	0.030	0.089	0.736	0.549
female: mother home, father intermediate	0.076	0.066	0.248	0.426
female: mother home, father professor	0.159	0.099	0.111	0.200
male: mother teacher, father unskilled worker	-0.036	0.094	0.702	0.621
male: mother teacher, father intermediate	0.132	0.071	0.063	0.226
male: mother teacher, father professor	0.087	0.090	0.337	0.426
male: mother home, father unskilled worker	0.142	0.096	0.140	0.217
male: mother home, father intermediate	0.058	0.069	0.404	0.474
male: mother home, father professor	0.066	0.092	0.474	0.474
number of observations	1248			

Note: estimates of (differences in) callback rates for September 2018, without control variables. ‘est’ provides the callback rate for the group ‘female: mother teacher, father intermediate’, as well as the differences in callback rates of all other groups relative to ‘female: mother teacher, father intermediate’. ‘boot se’ reports bootstrap standard errors clustered at the employer level. ‘raw p-val’ gives the p-values not accounting for multiple hypothesis testing. ‘adj p-val’ provides adjusted p-values accounting for multiple hypothesis testing.