ONLINE APPENDIX FOR:

"The Gender Application Gap: Do men and women apply for the same jobs?"

by Jonas Fluchtmann, Anita M. Glenny, Nikolaj A. Harmon and Jonas Maibom

A Online Appendix: Data and measurements

A.1 Data sets and sample selection steps

Our sample consists of UI recipients of Danish nationality entering new UI spells from September 2015 to September 2017. A new UI spell is defined when an individual who has not received UI benefits in the previous 4 weeks is observed with at least 4 consecutive weeks of UI payouts. For each UI recipient we identify and add all submitted job applications that have been registered in the Joblog system during the unemployment spell. The raw Joblog data is organized in several different databases with information on each edit (and save) of a given application entry. To generate the Joblog data used in this paper, we therefore pre-process these data sources and only include applications that were actually sent, and we further only select the first version of a given Joblog entry.[38]

To construct our final data set we make some additional sample restrictions, the effect of which we show in Table A.1 and also discussed in the main text. First, from the sample of new UI spells, we only consider the UI spells lasting at least 8 weeks. Second, we restrict our sample to individuals who register at least 4 applications in Joblog during their respective unemployment spell. Note that this condition is effective after dropping the last 4 weeks of applications, the individuals therefore need to have at least 4 registered applications during their UI spell. Third, we restrict our sample to individuals who leave UI for employment within the first year of their unemployment spell. Fourth, we drop all applications made in the last four weeks before entering employment. The data shows a drop in the number of applications that people register in Joblog about one month before they

[38]Besides documenting search activity to qualify for UI, the Joblog section of the Jobnet website was developed with the goal of helping job seekers keep track of their job search. In addition to submitting information on jobs that the worker has applied for, workers can also use the Joblog form to register and keep track of vacancies that the worker is considering applying for in the future and to register other job search events such as being called for an interview or being rejected. However, the coverage of these other events is much lower, and we only use data on the formal job application that UI recipients report that they have applied for.

enter employment, reflecting that individuals have already accepted their new job at this point and are just waiting for it to start. We therefore drop applications from the last four weeks before the new jobs start based on the median transition time between a successful application and starting a job, as applications made while waiting for the new job to start may not represent an individual's general application behavior.[39]

To the job application data we add data from three administrative data bases: These data sources are IDA, BFL and DREAM. IDA, the Integrated Database for Labor Market Research, is a matched employer-employee panel containing socioeconomic information on the entire Danish population. BFL, the Employment Statistics for Employees, contains monthly data on jobs, paid hours of work and earnings.[40] DREAM, is an event-history data set created by the ministry of employment tracing the participation of individuals in public income support programs at a weekly level. All data sets are available through servers at Statistics Denmark (see `https://www.dst.dk/en/TilSalg/Forskningsservice`). We link applied-for firms to firms in the BFL registers using a sting matching procedure which we explain in Section A.2.

Our final data set allows us to examine the characteristics of the firms and jobs that all male and female UI recipients apply to and compare them to the characteristics of the firms and jobs they eventually end up in.[41] For some of our measures of job characteristics we have missing values for some of the applications, in Tables A.2 we document the prevalence.[42] In the analysis we simply leave out job applications when the job characteristics in question are missing.

[39]The median transition time is 6 weeks, but we assume it takes 2 weeks from the application to the eventual job offer. The law requires active job search until the actual UI benefits stop (and the job starts), but naturally the type of jobs applied to are different, and the UI funds explicitly write that permanent positions should not be applied to when future employment is already secured. From 2019 the requirement was changed such that individuals with less than 6 weeks until the beginning of a new job will no longer have to register applications.

[40]The hourly wage measure we use is based on recorded monthly earnings divided by the recorded monthly hours in the job in the first full month of employment. This measure has the highest coverage in our sample as it is available already after one month of employment. The gender wage gap based on e.g. an average over 4 months of earnings is very similar, and the gap is likewise similar when we exclude observations where hours worked have been imputed by Statistics Denmark, see `https://www.dst.dk/da/Statistik/dokumentation/Times/beskaeftigelse-for-loenmodtagere/ajo-loentimer`.

[41]The occupation and industry classifications are available to several degrees of detail, grouped in major, sub-major and minor groups. Occupations are based on the Danish version of the ISCO classifcation (DISCO) and are grouped with 9, 55 or 153 respective occupations (referred to below as 1-, 2- or 3-digit groupings). The industries are based on NACE Rev. 2 and are grouped in 10, 21 or 38 respective industries.

[42]The typical reason for missing job characteristics is that we were unable to link the application to either the firm id or the specific occupation, see Section A.2. The higher shares of missing applications for the industry and firm wage levels (AKM firm fixed effects) reflects that either the firm match was unsuccessful or the firm is so new (small) that e.g. the industry affiliation is not recorded in the employment register (BFL) or it is not a part of the connected set. See also Section A.3.

In Figure A.1 we plot the survivor function for a version of our main analysis sample where we do not require individuals to find employment within 52 weeks. We also plot the average number of registered applications for each week in unemployment for the main analysis sample. As we discuss in Section 3.3 the average weekly number of applications during the unemployment spell is around 1.5 applications. Finally, in Figure A.2 we report the distribution of the number of submitted applications per week in our initial and final sample. Note that we discuss the dynamics in job search further in Section C.1.

A.2 Data matching algorithm

Before matching reported job titles and firms to official classifications and registers, we perform an extensive cleaning of these entries. In this step, we streamline the notation between source and target files and correct obvious spelling mistakes.

As a first step in the actual matching, we use the self-reported job titles and link these to the official Danish occupational codes (DISCO). We exploit that many of the self-reported job titles have the actual occupation as a part of the self-reported title. Thus, as a first step we identify occurrences of the DISCO occupations in the reported job titles. We only consider as 1:1 matches in this step (43.4 percent), i.e. if a certain job title links to several occupations we do not treat it as a match. For remaining unmatched entries, we manually match some job titles to occupations as many job titles use acronyms that do not match to the standard classification.[43] This adds about 27.2 percent to the matches. Finally, we also use some fuzzy matching techniques on the remaining unmatched observations to circumvent misspellings in the job titles, adding the manual titles from the step before. We rank the potential matches along several scoring functions and only pick consistently high-ranked matches. For this we use compget, speedist and soundex routines from SAS as well as sub-string occurrences which adds 10.9 percent. Overall, we can thus map 81.5 percent of the applications to a DISCO group.

As the second matching step, we link the reported firm information to firm identifiers. With the mandatory reporting of firm name, zip code and city, we develop a matching procedure which matches this information to the official firm registers recording all Danish firms (CVR-register).[44]

[43] For example 'social og sundhedshjælper', Danish for social and health care workers, are most often reported as 'sosu-hjælper'.

[44] The Danish central firm register (CVR-register in short) contains information on companies officially registered

We can then use these links to identify firms in the registers at Statistics Denmark (BFL). Our matching procedure on firms also starts with perfect matches, using both firm name and zip codes. Here we have a 1:1 match for 66.3 percent of the applications in Joblog. We further add the sub-string matches which are spatially the closest to the reported firm address (13.9 percent). To link applications which we cannot match exactly on firm names, we employ a fuzzy matching procedure using the matchit command in STATA to identify the 50 closest matches. We then test these 50 potential matches using several scoring functions besides the one obtained from matchit. For each of the scores (5 in total), we calculate the ranking of the 50 potential matches (rank 1 is the best) and identify the "correct" match as the match which receives the best average rank (the scores we use are Bi-gram Similscore, Token, TokenSound from matchit and the compget and speedist functions in SAS). This adds further 6.2 percent to the matches, so we end up with an overall firm match of 86.4 percent.

A.3 Measuring firm wage levels using AKM firm fixed effects

We use our matched employer-employee data to estimate an AKM model (Abowd et al., 1999) and use the estimated firm fixed effects as a measure of the firms wage premium. The AKM model captures implied firm fixed effects on wages, i.e. the firms' wage premium, by identifying moves of workers from one firm to another while simultaneously absorbing individual wage components in worker fixed effects. The separate identification of firm fixed effects relies on the connection between firms in terms of worker movements. Thus, the firm fixed effects can only be recovered for the set of connected firms. We take advantage of the rich administrative data on the whole Danish working population, in particular the BFL data set (see Section A.1) covering monthly salaries, to construct a matched employer-employee panel based on earnings in March and September within a given year. We focus on the years 2008 to 2017 and arrive at 265,425 (connected) firms for whom we get estimates of firm effects (this amounts to around 95 percent of all firms present in the data in the sampling period).

Our estimation uses hourly wages and in addition to worker and firm effects we also include calendar month fixed effects to absorb any aggregate time trends. The hourly wage measure we use

in Denmark. The register covers all firms, with the exception of privately held companies with an annual turnover below 50,000 DKK (about 7,500 USD). Each firm is registered with a uniquely identifiable CVR number that's linkable to Danish administrative data sets.

is based on recorded monthly earnings divided by the recorded monthly hours. After the estimation, we subtract industry specific averages of firm effects from the estimated firm effects, and divide through by industry specific standard deviations, to ensure that the rankings we obtain account for industry differences in the distribution of firm effects. Further, to guarantee equal size of the decile bins, we employment weight these rankings with the number of employees in each firm as of August 2015, the month before we observe applications in Joblog.[45] For completeness we also report gender gaps based on the non-standardised firm effects when relevant.[46]

A.4 Constructing typical wages

To construct our measure of the typical wage paid in a job with certain characteristics, we use data on the new jobs in our analysis sample to estimate a model that predicts wages based on the characteristics of the job. Since this is a pure prediction problem, we use a LASSO-based machine learning approach. Specifically, we consider a linear regression with log wages as the outcome variable and a very large number of potential explanatory variables based on the available job characteristics in our data. We then use LASSO estimation to select the subset of these variables that most efficiently trades off predictive power in-sample against the risk of overfitting.

As the baseline set of explanatory variables we include dummies for the industry and occupation of the job at both the 1-, 2- and 3- digit levels and the within-industry-standardized AKM firm fixed effect.[47] To handle firm's where no AKM fixed effect can be estimated, we normalize their AKM fixed effect to zero and include a dummy for the fixed effect being missing. Finally, we include all pairwise interactions between the variables for a total of 10,407 baseline explanatory variables. We rely on the Rigorous-LASSO approach of Belloni et al. (2016) to choose the regularization parameters for the LASSO estimation. Because some individuals show up with several UI spells in our data, we allow for clustered disturbances at the individual level in estimation. The estimation

[45] This additionally implies that we ensure that our results are not driven by estimated firm fixed effects from smaller firms which are known to be imprecisely estimated, see e.g. Andrews et al. (2008).

[46] As a robustness check we have also repeated the above steps separately by gender to obtain gender specific firms effects. For these results and further discussion, see Section B.6.2.

[47] Note that the resulting set of dummies thus exhibit perfect multicollinearity by definition. Because of the penalization term in the LASSO objective function, however, perfect does not create a problem for estimation. Including all the dummies, however, is advantageous because the LASSO aims to select a parsimonious model with high predictive power. An obvious candidate for such a model is one that includes dummies for most of the aggregate 1-digit categories but in addition includes dummies for some 2- or 3-digit categories. The final prediction model indeed has this flavor (see Table A.3).

was conducted using the LASSOPACK implementation of Ahrens et al. (2019a,b).

Out of the 10,407 baseline explanatory variables, the Rigorous-LASSO selects 233 variables. As the final step, we run a standard OLS regression with log wage as the outcome variable and these 233 variables as explanatory variables (so-called Post-LASSO OLS) to arrive at our final prediction model.[48] Table A.3 summarizes the estimation and final model.

In Table A.4, we examine the performance of our prediction model. In addition, to validating the use of the prediction model for our analysis, this allows us to draw lessons for future analyses on related data; the need to infer wages from job characteristics arises often when analyzing data on vacancies and/or applied-for jobs (see e.g. Marinescu and Skandalis (2021). To validate model performance we compute an out-of-sample R^2 based on a standard sample-splitting procedure: We randomly mark 80 percent of the new hires data as a training subsample. We then use only this subsample to estimate (train) the model. Finally we use the model to predict for the remaining 20 percent of the data and compute the R^2 for these predictions.

The first row of Table A.4 show that our main LASSO prediction model has an out-of-sample R^2 of 0.201. As a natural benchmark, the second row shows results for a standard OLS approach that simply regresses log wages on the standardized AKM fixed effects as well as industry and occupation dummies at the 3 digit level. Unsurprisingly, this OLS performs reasonably well with an out-of-sample R^2 of 0.192, however, as expected our main LASSO-based approach performs noticeably better. This illustrates the advantage of using a modern Machine Learning approach to prediction. Finally, the last two columns show how much predictive power is lost if the AKM fixed effect is not used or if all firm information is excluded and only occupation information is used. This is informative for settings where linked employer-employee information is not available or where no firm information is available at all. As expected the AKM fixed effect is very important for the models predictive power; dropping just this variable cuts the out-of-sample R^2 down to 0.164. We note, however, that a LASSO model based only on occupations still retains an R^2 of 0.138. In line with previous results relating wages to occupations and job titles, meaningful information about wages can be learned simply from the occupation of a vacant job (Marinescu and Wolthoff (2019)).

After estimating our main LASSO prediction model, we use it to compute our measures of the

[48] As is common, some of the 233 variables selected by the Rigorous-LASSO turn out to be collinear and thus drop out in the OLS regression. As a result, the total number of variables in the final OLS regression is 117.

typical wage for new jobs as well as for applied-for job in our application data by simply predicting the log wage from the job's characteristics. Since some applied-for jobs in our application data cannot be linked with firm/or occupation information (see Section A.2), we estimate two alternative prediction models by applying exactly the same procedure as above but excluding either firm or occupation information from the baseline set of explanatory variables. We use these alternative models to fill in typical wages also for the applications that cannot be linked with firm/or occupation information.

We note that the prediction models we use to compute typical wages only ever include job and firm characteristics but never include any worker characteristics. Our measure of the typical wage in a given job thus makes no attempt to capture that workers with different characteristics might face different wages in the same job.[49]Throughout our analysis of gender differences in applications and hiring outcomes we condition on observables so that we are in fact comparing men and women with similar labor market observables. A particular issue arises, however, if the typical wage offered in a particular type of job depends directly on gender (see Section 4.4). As a robustness check we therefore also present results where our measure of typical wages in a given job is based only on the wage paid to either men or women (See Appendix B.6). We do this simply by redoing the final Post-LASSO OLS estimation on either the male or female half of the sample.

In unreported results, we have also experimented with using alternative prediction approaches to the construction of our typical wage measure, including other machine learning approaches or simple linear regressions with a smaller set of variables. Our results are not sensitive to using these other approaches.

A.5 Selecting our set of conditioning variables

To discipline which labor market observables we condition on we follow recent suggestions in the literature (e.g. Angrist and Frandsen, 2022; Athey and Imbens, 2019; Mullainathan and Spiess, 2017) and rely on a Machine Learning procedure. Specifically we use the double-LASSO procedure of Belloni et al. (2014) to select the most important variables for explaining the gender wage gap. We start by specifying a very large baseline set of variables that ex ante could be important to

[49]Formally, our measure of typical wage in a given job reflects the wage paid to the average individual that is hired into this type of job in the data.

condition on when analyzing the gender wage gap. Using data on all the individuals in our analysis sample, along with the wages in their new jobs, the double-LASSO procedure then involves two separate LASSO regressions: First the LASSO is applied to a regression that has log wage in the new job as the outcome variable and includes the full set of baseline variables as regressors. Intuitively, this step selects out any variables in the baseline set that are relevant predictors of wages. Second, the LASSO is applied to a regression that has a female dummy as the outcome variable but again includes the full set of baseline variables as regressors. This step selects out variables that are significant correlates of gender. Combining all the variables selected in each of these two steps then gives the set of most important variables for explaining the gender gap in wages.[50] In each of the steps in the double-LASSO, a data-driven procedure is used to determine the penalty parameter for the LASSO estimation. See Belloni et al. (2014) and Urminsky et al. (2016) for additional discussion and formal results.

The baseline set of potential variables that we include in the two regressions consists of a set of 4,196 variables: To capture educational differences, the set contains years of education, as well as dummies for the field of study. To capture additional differences in general human capital, the set includes age, total work experience and total work experience over the last five years. To capture additional differences in specific human capital, the set includes dummies for the sector, industry and occupation of the previous job as well as continuous measures for the total work experience over the last five years in each of the different industries and occupations. To capture differences in dependence on public transfers, the set includes the total time spent on unemployment insurance, social assistance and other public transfers over the last five years. When including dummies for industry, occupation or field of education and when including continuous measures of industry or occupation-specific experience, we always include all possible measures at both the 1-, 2- and 3-digit level.[51] Finally, all variables are also interacted with both age, years of education, total work

[50] The intuition here is that in order to play a significant role in explaining gender differences in wages a variable has to be strongly correlated with either wages or gender. Variables that are weakly correlated with both however should not play an important role in explaining gender differences in wages. These are exactly the variables excluded by the double-LASSO.

[51] Our coding of occupations, industries and fields of education are based on the official definitions by Statistics Denmark, see also footnote 41. We note that the resulting set of included dummies exhibit perfect multicollinearity by definition. Because of the penalization term involved in the LASSO objective function, however, perfect multicollinearity does not create a problem for estimation. Including all the dummies, however, is advantageous because the LASSO aims to select a parsimonious set of variables with high predictive power. An obvious candidate for such a a set is one that includes dummies for most of the aggregate 1-digit categories but in addition includes dummies for some 2- or 3-digit categories. The final set of selected variables indeed has this feature (see Table A.5).

experience and work experience over the last five years.[52]

The double-LASSO selects 332 of these variables which we use as our observable characteristics to condition on throughout the main analysis.[53] The estimation was carried out using the PDSLASSO implementation of Ahrens et al. (2018). Table A.5 summarizes the final selected set of variables that we condition on.

A.6 Propensity score reweighting for descriptive results

As discussed in Section 3.4 and Section A.5, we use propensity score reweighting in all of our analysis. Using the notation introduced in Section 3.4, the reweighting scheme involves reweighting woman i by a weight equal to $\frac{\hat{p_i}}{1-\hat{p_i}}$, where $\hat{p_i}$ is an estimate of the conditional probability of being male given observables, $P(m_i = 1|x_i)$.

After selecting the set of variables to include in our vector of observables, x_i, we follow the standard in the literature and estimate a logit model for the probability of being male, using the variables in x_i as our explanatory variables. We then obtain the $\hat{p_i}$s as the predicted probabilities from this model and use these to reweight the women in our sample.

In Figure A.3 we show the distribution of the estimated propensity scores in our sample. Note that we trim our sample to avoid very small or very large weights, see Table A.1. Specififcally we trim all observations with an estimated propensity score larger than 0.99 or smaller than 0.01.

A.7 Decomposition additional details

In this section we briefly translate the main insights and methodology of DiNardo et al. 1996 and Fortin et al. 2011, showing how propensity score reweighting can be used to construct estimates of the counterfactual hiring probabilities $P_{A,X}^{W\tilde{}}(y)$ and $P_X^{\tilde{W}}(y)$ underlying our decomposition exercise and introduced in Section 5.

We start by considering the counterfactual hiring probability for women if they had the same distribution of observables as men.:

[52]In implementing this, we allow for variables to be interacted with themselves so that our baseline set includes squared terms in age, years of education, total work experience and work experience over the last five years.

[53]As is common, some of the 332 variables selected by the Rigorous-LASSO turn out to be collinear and thus in practice drop out from our conditioning set when we implement our propensity weighting procedure (see Section A.6). As a result, the total number of variables in the model that we use to estimate propensity scores is 302.

$$P_X^{\tilde{W}}(y) = \iint P^W(y|a,x) f_{a|x}^W(a|x) f_x^M(x)\,\mathrm{d}a\,\mathrm{d}x$$

Multiplying and dividing by $f_x^W(x)$ inside the integral, we can rewrite this as follows:

$$P_X^{\tilde{W}}(y) = \iint P^W(y|a,x) f_{a|x}^W(a|x) f_x^W(x) \Psi_X(x)\,\mathrm{d}a\,\mathrm{d}x$$

Here we have defined $\Psi_X(x) = \frac{f_x^M(x)}{f_x^W(x)}$. The first insight is that $P_X^{\tilde{W}}(y)$ is simply a weighted expectation of $P^W(y|a,x)$ over the set of all women weighted by $\Psi_X(x)$:

$$P_X^{\tilde{W}}(y) = E\left[\Psi_X(x) P^W(y|a,x)|m=0\right]$$

It follows that if the weighting function $\Psi_X(x)$ was known, $P_X^W(y)$ could be estimated by applying the weighting function and then simply computing the share of women hired into job type j in the weighted sample.[54] Now, by an application of Bayes rule, $\Psi_X(x)$ is proportional to a simple function of the conditional probability for being male conditional on observable characteristics x (the propensity score):

$$\Psi_X(x) \propto \frac{P(m=1|x)}{1-P(m=1|x)}$$

It follows that$P_X^{\tilde{W}}(y)$ can be estimated via propensity score reweigthing. We estimate a logit model for the likelihood of being male as a function of our observable characteristics x and then use the predicted probabilities from this model to reweight the women in our sample before computing hiring probabilities. As noted in the main text, this is equivalent to the propensity score reweighting we use to condition out observable differences between men and women, as introduced in Section 3.4.

Next consider the counterfactual hiring probability that women would have faced if they also

[54]To see this more clearly, let $I(y)$ be an indicator for ending one's UI spell by being hired into job type y and note that we have:

$$\begin{aligned} P_X^{\tilde{W}}(y) &= E\left[\Psi_X(x) P^W(y|a,x)|m=0\right] = E\left[\Psi_X(x) E\left[I(y)|a,x\right]|m=0\right] \\ &= E\left[E\left[\Psi_X(x) I(y)|a,x\right]|m=0\right] = E\left[\Psi_X(x) I(y)|m=0\right] \end{aligned}$$

The direct empirical counterpart of the last expectation is then the share of women hired into job type y after applying the $\Psi_X(x)$ weights: $\frac{1}{N_W}\sum_{m_i=0} I(y_i)\Psi_X(x_i)$ (here subscript i refers to individuals in the data, and N_W is the total number of women in the data).

had the same distribution of application behavior as men:

$$P^{\tilde{W}}_{A,X}(y) = \iint P^W(y|a,x) f^M_{a|x}(a|x) f^M_x(x) \, \mathrm{d}a \, \mathrm{d}x$$

Letting $f^M_{a,x}$ and $f^M_{a,x}$ denote the joint distribution of application behavior and observables for men and women respectively, we rewrite this in a similar way as before:

$$P^{\tilde{W}}_{A,X}(y) = \iint P^W(y|a,x) f^W_{a,x}(a,x) \Psi_{A,X}(a,x) \, \mathrm{d}a \, \mathrm{d}x$$

Here we have defined $\Psi_{A,X}(a,x) = \frac{f^M_{a,x}(a,x)}{f^W_{a,x}(a,x)}$. Similar to before we see that this implies that the counterfactual hiring probability can be estimated by reweighting the women according to the weighting function $\Psi_{A,X}(a,x)$. Also as before, an application of Bayes rule shows that the weighting function is proportional to a simple function of the conditional probability for being male conditional on both observable characteristics x and application behavior a (a different propensity score):

$$\Psi_{A,X}(a,x) \propto \frac{P(m=1|a,x)}{1-P(m=1|a,x)}$$

It follows that$P^{\tilde{W}}_X(y)$ can also be estimated using propensity score reweigthing. We estimate a logit model for the likelihood of being male as a function of our observable characteristics x and application behavior and then use the predicted probabilities from this model to reweight the women in our sample before computing hiring probabilities.

A.8 Additional details on non-wage characteristics

A.8.1 Measuring family-friendliness

To construct a simple measure of how family-friendly a firm is, we use data on how much parental leave employees at the firm tend to take when they become parents. The basic idea here is that employees will tend to take longer leave if their firms offers more generous parental leave terms and/or are very tolerant towards employees going on leave.[55] Family-friendly firms that offer

[55] During our sample window, government-mandated parental leave rules are as follows: Mothers are entitled to the following weeks on leave with compensation by the government: 4 weeks of leave just before birth, 14 weeks of maternity leave post birth, and subsequently 32 weeks of parental leave which can also be used by the father. Fathers are further entitled to 2 weeks of paternity leave immediately after birth. The government compensation during leave is at the UI benefit level, however most employment contracts in Denmark offer periods of leave where the

generous leave packages and are supportive of employees leave-taking should thus see longer parental leave periods among their employees.

We extract information on the duration of leave through the Danish register on sickness benefit claims (SGDP) which contains all information about benefits paid out by the government in connection with sickness and childbirths. Because this data covers both reimbursements made to firms with workers on leave and payments made directly to the workers on leave, it can be used to infer how much parental leave an employee takes. We focus on benefit claims after 2011 for both men and women and select all payouts which are related to childbirth and where the worker eventually returns to the same pre-birth employer. We then accumulate days with payments within individuals. As our data does not come with readily available information on birthdays for children, we infer these from starting a new parental leave spell (or having more than half a year between payments).[56] Finally, we calculate the average days on leave per birth at the firm level.

In Figure A.4 we plot the distribution of average leave lengths at the firm level. Before computing the the average leave taken at each firm, we correct for the gender of the employee to account for the fact that women take much longer leaves than men. We do this by simply demeaning the leave periods by gender.

Obviously, the average parental leave length will be undefined for firms that do not experience an employee or their partner giving birth at any point during the time period we consider. Since such firms may be likely to be at the bottom of the distribution in terms of family-friendliness,[57] however, we do not exclude them from the analysis but include "no birth" as a separate category when analyzing family-friendliness.

A.8.2 Non-wage characteristics and wage correlations

In Table A.6 we show the result from simple linear regressions where the dependent variable is the typical wage associated with a given application (see Section 3.2) and the independent variables are

worker receives additional compensation by the firm. The duration of these employer coverage periods differ by firms, occupations and collective agreements. During the employer coverage periods the government subsidy is instead paid to the firm, and the firm simply continues to pay the worker his/her wage. When the employer coverage period ends the government payment is redirected and paid directly to the worker.

[56]Focusing on mothers this corresponds well with the official birth statistics. Note that some women may transition from one leave to another, thus we censor the length of leave at 14 months and regard any subsequent leave as a new birth/spell. This is however very limited in our data.

[57]This type of firm likely also include rather new firms or firms which are so small that births are rather infrequent.

our selected non-wage characteristics.[58] The sample contains all the the submitted applications, and we cluster standard errors at the level of the unemployment spell. The estimates show that part-time jobs, jobs involving shorter commutes and jobs at family-friendly firms all tend to offer lower typical wages in our data. As a result, if women send more applications to jobs characterized by e.g. shorter commute compared to males, this should translate into gender application gaps suggesting that women to a larger extent apply to jobs with lower typical wages.

A.9 Further documentation of application behavior and data

A.9.1 Survey about Joblog usage

Table A.7 and A.8 present results from a survey conducted among Danish UI recipients by Mahlstedt et al. (2019) in March 2018.[59] Table A.7 reports survey answers about how individuals log applications in Joblog. Looking at the first column, 41 percent of respondents report that they always log all the jobs they have applied for in Joblog regardless of whether they have fulfilled the logging requirements. An additional 21 percent report that they only log applications up to the point where they have satisfied their logging requirements but that they rarely apply for more jobs than what is required. Putting these together suggest that Joblog has close to full coverage for 63 percent of respondents. For the remaining 37 percent, however, the survey responses suggest that the Joblog data often misses some job applications that they have made beyond the required number. The second and third columns report corresponding numbers by gender. These show a similar pattern overall, although men are somewhat more likely to say that they often apply to jobs that they do not register. 42 percent of men say that they often apply to more jobs than they register, while only 32 percent of women say so.

To get a sense of how many applications may be missed by the Joblog data, Table A.8 presents survey responses about the total number of job applications sent the past month and the number of job applications sent that were not registered in Joblog. In addition, the Table also shows the actual number of registered Joblog applications made by the survey respondents in the month before the survey. This was computed by linking survey responses with the actual Joblog data. On average, survey respondents report applying for 11.5 jobs in total over the past month. Of those

[58]We remove applications where the relevant job characteristics are missing from the analysis when necessary.
[59]We thank the authors for making this data available.

jobs, survey respondents on average say they failed to register 2.4 jobs in Joblog. This suggest that Joblog covers 80 percent of actual applications. The bottom of the table instead shows that average number of jobs respondents actually registered in Joblog was 8.0. Relative to the total number of reported applications, this suggest that respondents on average failed to register 3.5 applications, implying that Joblog on average covers 69 percent of all applications.[60] The table also reports separate numbers for men and women. Women self-report failing to register 2.1 jobs on average, implying that joblog covers 82 percent of applied-for jobs for women. Alternatively, comparing total reported applications to actual joblog registrations, suggest that women on average fail to register 3.8 applications, corresponding to a coverage of 68 percent. Corresponding calculations for men suggest that Joblog covers between 72 and 76 percent of applied-for jobs for men.

A.9.2 Coverage and representativeness of the Joblog data

In this section we present additional evidence and validity checks regarding the quality of the Joblog application data. In doing so we exploit the fact that the data is linked to actual job outcomes.

First, we verify how the application data relates to actual hiring outcomes. If the Joblog application data accurately capture actual application behavior, we would expect the application data to be highly predictive of the type of job that each UI recipient ends up being hired into. To assess whether this is the case, we benchmark the predictive value of the application data against a known strong predictor of job outcomes: the characteristics of UI recipients previous job.

Table A.9 compares how the Joblog application data and prior job characteristics predict respectively the industry, the occupation, the firm wage level or the typical wage level of a UI recipient's new job. Each column of the table corresponds to a different prediction model estimated on our analysis sample. When predicting the industry of the new job, we use a simple multinomial logit model that includes either dummies for the industry of the previous job, the share of applications going to each industry or both sets of variables. Similarly, when predicting the occupation of the new job we use a multinomial logit model that includes either dummies for the occupation of

[60] This difference could reflect imperfect recall among survey respondents or could relate to measurement error from the timing of registered jobs and/or the precise interpretation of the survey question. Registering applied-for jobs in Joblog can be done retroactively so the interpretation of the survey question could either refer to the date at which applications were sent or to the date at which the application was entered into the Joblog system. Additionally, the fact that UI recipients are able to register other activities besides formal job applications introduces some ambiguity about the interpretation of the survey question (if for example UI recipients have registered that they reached out to a friend about a specific job).

the previous job or the share of applications going to each occupation. For both industries and occupations, we exclude one small industry/occupation for which the model obtains near perfect predictions for a few observations.[61] When predicting the firm wage level or the typical wage of the new job, we use a simple linear regression that includes either the firm wage level or the typical wage of the previous job, or includes the mean of the firm wage level or typical wage across the applied-for jobs. For the linear regression models, we measure the predictive power simply using the regression R^2. For the multinomial logit models, we use McFadden's *pseudo*-R^2.

Looking across Table A.9, we see that models that predict job outcomes only using application data perform quite similarly to models that instead use prior job characteristics. Application data does do markedly worse than prior job characteristics when predicting the occupation of the new job (column (4) vs (5)) but only slightly worse for firm wage level and the typical wage (columns (7) and (10) vs. (8) and (11)). At the same time, application data actually does better than prior job characteristics when predicting the industry of the new job (column (1) vs. (2)).

For models that include both prior job characteristics and application data (columns (3), (6), (9) and (12)), we see that application data remains highly predictive even after prior job characteristics have been conditioned on. Adding the application variables alongside prior job characteristics always leads to sizeable increases in the (*pseudo*-)R^2 relative to models that only use prior job characteristics. Moreover the application variables are always highly statistically significant in the combined models. Overall, we conclude that the Joblog application data is highly predictive of later job outcomes.

In Tables, A.10 and A.11, we repeat the prediction exercise separately for the men and women in the data. Results are very similar and we see little indication of systematic differences in the predictiveness of job applications for later job outcomes. While job application information appears to predict occupational outcomes slightly better for men, they instead appear to predict wage outcomes slightly better for women. Moreover, the differences are small throughout.

As a further check on the representativeness and quality of the Joblog application data, we examine how often we are able to trace a new hire back to a job application that is contained in

[61] Specifically, we exclude individuals from the sample who find a job in the smallest industry or occupation, respectively, as well as individuals whose prior job was in this industry or occupation. Results are almost identical if these observations are included, however, we see indications that the likelihood function becomes ill-behaved in some specificaions in this case, reflecting that some observations are predicted nearly perfectly.

our data. For 47 percent of the new hires, we are able to identify a previous application that the UI recipient sent to the firm in question. This is informative about the representativeness of the data. To see why, assume that the Joblog data covers a share r of all applications and that the share of applications that we successfully match to firms in our data matching procedure is s. In this case, our data will contain firm information for a share $s \cdot r$ of all applied-for jobs. Next assume that the fraction of jobs that stem from a job application is j. If the applied-for jobs in our data are a representative subset of all applied-for jobs, the share of new hires that we should be able to trace back to an application, t, should then be:

$$t = j \cdot r \cdot s$$

Based on independent survey data from Table A.8 we estimated that the raw Joblog data contain between 69 and 80 percent of all applied-for jobs, that is r is between 0.69 and 0.80. Furthermore, as described in Section 3.1, $s = 0.86$ in our data matching procedure. Finally, to gauge the share of hires that stem from a job application, j, we rely on Statistics Denmark's official survey *Arbejdskraftsundersøgelsen* on how unemployed Danes report landing their first job out of unemployment (Engman (2019)). In these data, 11 percent of respondents report landing their job in a way that is very unlikely to have involved the worker applying for the job (the job resulted from work at a temp agency, they got the job via their educational institution as an internship or the job seekers themselves advertised publicly), while 58 percent of respondents report landing their job in a way that almost surely involved making a formal job application (they themselves applied to a posted position, they applied to a firm with no posted positions or they were directed to the job by the employment agency or other authorities).[62] To arrive at an estimate for the fraction of hires that stem from workers applying for the job, we simply assume that half of the remaining jobs involved a job application. This implies that about 73 percent of new hires out of unemployment involve the worker applying for the job at some point so that $j = 0.73$.[63]

Plugging in these values, we see that if the applied-for jobs in our data is representative, the

[62]The remaining respondents report landing their jobs through channels that may have involved applying for the job application but may also have involved receiving a job offer more directly. This includes finding the job through an acquaintance or finding a job after having been contacted by the firm.

[63]Alternatively, we could use 0.58 as a lower bound on r and use 0.89 as an upper bound. Plugging into the formulate above, we then see that if the applied-for jobs in our data is representative, the share of new hires that we should be able to match, t, should be between $0.58 \cdot 0.69 \cdot 0.86 = 0.34$ and $0.89 \cdot 0.80 \cdot 0.86 = 0.61$.

share of new hires that we should be able to match, t, should be between $0.73 \cdot 0.68 \cdot 0.86 = 0.43$ and $0.73 \cdot 0.80 \cdot 0.86 = 0.50$. As noted, we in fact have $t = 0.47$ in our data, consistent with the data containing a representative subset of all applied-for jobs.

Finally, we can also compute how often we are able to trace a new hire back to a job application that is contained in our data for men and women respectively. For 41 percent of all new male hires, we are able to identify a previous application that the UI recipient sent to the firm in question. For women, the corresponding number is 53 percent. A likely reason for this difference is that Danish men are more likely to find jobs in ways that do not involve a formal job application (see e.g. Engman (2019) for evidence of this for the overall labor market). This mechanically implies that we should be able to match a smaller share of new male hires to a job application in our data. Because we do not have reliable data on how often this occurs for unemployed men and women, however, we cannot assess quantitatively how the gender-specifc match rates match up with our other data.

Table A.1: Sample selection and trimming

	Individuals	Spells	Applications
Inflow	**227.515**	**261.529**	**7.422.148**
Minimum 8 weeks spell length	177.145	194.660	7.019.513
Spells with > 4 applications	170.304	185.959	7.008.284
Employment within 52 weeks	105.879	114.375	3.439.690
Censoring last 4 weeks of applications	105.879	114.375	2.911.585
Analysis sample	**105.879**	**114.375**	**2.911.585**
After trimming for descriptive analysis†	100.268	108.173	2.790.250
After trimming for decomposition ††	93.550	100.504	2.631.684

Notes: The top of the table shows the number of individuals, unemployment spells as well as number of applications in the base UI inflow data and after applying each of our four sample restrictions. The bottom of the table shows how the analysis sample changes when we apply trimming to remove observations with extreme propensity scores. † refers to the trimming we use in our descriptive analysis where propensity scores are estimated based only on labor market observables (see Section 3.4). †† refers to the additional trimming we use in the decomposition exercise where propensity scores are estimated based on both labor market observables and application behavior (see Section 5).

Table A.2: Share of missing job characteristic

	Firm ID	**Occupation**	**Industry**	**Firm Wage Level**	**Typical Wage**	**Wage**
Application	0,185	0,193	0,329	0,330	0,068	1,000
Hires	0,001	0,083	0,059	0,060	0,000	0,102

Notes: The table shows the share of missing job characteristics both for applications and for the jobs that UI recipients are hired into in the analysis sample.

Table A.3: Prediction model summary

Model and estimation summary:	
Explanatory variables in baseline set:	10,407
Variables selected in Rigorous-LASSO:	233
Parameters in final Post-LASSO OLS model:	117
R^2 for Post-LASSO OLS model (in sample):	0.202
Summary of selected variables:	
Standardized firm fixed effect	
7 dummies for 1-digit occupations	
7 dummies for 2-digit occupations	
17 dummies for 3-digit occupations	
7 dummies for 1-digit industries	
7 dummies for 2-digit industries	
7 dummies for 3-digit occupations	
161 occupation-industry interactions	
9 occupation-firm fixed effect interactions	
10 industry-firm fixed effect interactions	

Notes: The table summarizes the main prediction model used to construct the measures of typical wages. The difference between the number of selected variables in the Rigorous LASSO and the number of parameters in the final Post-LASSO OLS model reflect that some of the selected variables are perfectly multicollinear (see footnote 48).

Table A.4: Prediction model performance

	Out-of-sample R^2	Estimator	Variables
Main LASSO model	0.201	Post-LASSO OLS (Rigorous penalty)	All occupation dummies (1, 2 and 3 digits), All industry dummies (1, 2 and 3 digits), AKM fixed effects + pairwise interactions
OLS benchmark	0.192	OLS	Occupation dummies, 3-digits Industry dummies, 3-digits AKM fixed effects
LASSO, no AKM fixed effects	0.164	Post-LASSO OLS (Rigorous penalty)	All occupation dummies (1, 2 and 3 digits), All industry dummies (1, 2 and 3 digits), + pairwise interactions
LASSO, occupation only	0.138	Post-LASSO OLS (Rigorous penalty)	All occupation dummies (1, 2 and 3 digits)

Notes: The table shows model performance for four different prediction models. The models are either based on OLS or on Post-LASSO OLS using the rigorous penalty approach of Belloni et al. (2016). The models differ also in terms of the set of baseline variables included, as shown in the last column. The out-of-sample R^2 of each model was computed based on an 80-20 train-test split of the data.

Table A.5: Summary of observables selected in double-LASSO

Non-interacted continous variables:

Age, years of education

2 continous measures of experience from 1-digit industries
9 continous measures of experience from 2-digit industries
12 continous measures of experience from 3-digit industries

4 continous measures of experience from 1-digit occupations
1 continous measures of experience from 2-digit occupations
32 continous measures of experience from 3-digit occupations

Non-interacted dummy variables:

4 dummies for the 1-digit occupation of most recent job
13 dummy for the 2-digit occupation of most recent job
29 dummies for the 3-digit occupation of most recent job

3 dummies for the 1-digit industry of most recent job
6 dummies for the 2-digit industry of most recent job
10 dummies for the 3-digit industry of most recent job

2 dummies for the sector of the most recent job
3 dummies for education field at the 1-digit level
8 dummies for education field at the 2-digit level
17 dummies for education field at the 3-digit level

Interactions involving only continous variables:

5 interactions involving only combinations of age, years of education or experience

2 interactions involving continous measures of past receipt of public transfers
47 interactions involving continous measures of experiences from specific occupations
10 interactions involving continous measures of experiences from specific industries

Interactions involving discrete variables:

42 interactions involving dummies for the occupation of most recent job
16 interactions involving dummies for the industry of most recent job
3 interactions involving dummies for the sector of the most recent job
50 interactions involving dummies for field of education

Notes: The table summarizes the set of 332 variables selected in the double-LASSO. Interaction terms always involve either age, years of education, total work experiene or work experience over the last five years as one of the two interacted variables. We also allow for variables to be interacted with themselves, corresponding to a squared term (see footnote 52).

Table A.6: Correlation between job characteristics and typical wages

	(1) Log typical wage	(2) Log typical wage	(3) Log typical wage	(4) Log typical wage
Commute	0.000240***			0.000227***
(minutes)	(4.78e-06)			(4.58e-06)
Part-time		-0.0712***		-0.0704***
(indicator)		(0.000584)		(0.000576)
Family-friendliness			-0.000143***	-0.000143***
(days of leave)			(3.70e-06)	(3.60e-06)
Constant	5.152***	5.179***	5.180***	5.173***
	(0.000474)	(0.000412)	(0.000445)	(0.000522)
Observations	2,638,100	2,638,073	2,638,100	2,638,073
R-squared	0.016	0.024	0.007	0.047

Notes: ***$p<0.01$, **$p<0.05$, *$p<0.1$. The table shows the result of regressing the log typical wages of an applied-for job in the main anlysis sample on different non-wage job characteristics. Jobs with missing data on any of the involved variables have been dropped in all specificatons. Standard errors are clustered at the spell level.

Table A.7: Survey question "Which of these statements best describes your use of Joblog?"

Answer:	**Share Overall:**	**Share Men:**	**Share Women:**
Fulfill requirements, often applied to more jobs	0.36	0.42	0.32
Fulfill requirements, rarely applied to more jobs	0.21	0.19	0.23
Always register all applied-for jobs	0.41	0.38	0.44
Never register applications	0.1	0.2	0.1
Number of respondents	1236	515	721

Notes: The table shows answers to the question ”Which of these statements best describes your use of Joblog?” based on the survey of UI recipients conducted in Mahlstedt et al. (2019).

Table A.8: Self-reported and registered applications in the previous month

	Mean, overall	Mean, men	Mean, women
Survey answers			
# of applied-for jobs	11.5	11.2	11.7
# of applied-for jobs not registered	2.4	2.7	2.1
Joblog data			
# of applied-for jobs	8.0	8.1	7.9

Notes: The to part of the table shows the reported number of job applications sent over the last month and the reported number of these jobs applications that were not registered in Joblog based on the based on the survey of UI recipients conducted in Mahlstedt et al. (2019). The bottom part of the table shows the actual number of jobs registered in Joblog by the survey respondents in the month prior to the survey.

Table A.9: Predicting job outcomes from application data vs. prior job characteristics

Job outcome:	Industry (1 digit)			Occupation (1 digit)			Firm wage level			Typical wage		
Model:	Multinomial logit			Multinomial logit			Linear regression			Linear regression		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Explanatory variables												
Characteristics of previous job	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Characteristics of applied-for jobs	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes
# of parameters	99	99	189	63	63	119	2	2	2	2	2	2
(pseudo-)R-squared	0,292	0,245	0,368	0,331	0,395	0,497	0,056	0,077	0,112	0,268	0,285	0,392
p-value, test of excluding applied-for job variables	< 0,01		< 0,01	< 0,01		< 0,01	< 0,01		< 0,01	< 0,01		< 0,01

Notes: The table examines the predictiveness of job applications and past job characteristics for the sample overall. Columns (1)-(3) correspond to multinomial logit models for the 1-digit industry of the UI recipients new job. Explanatory variables in these models are dummies for the 1-digit industry of the previous job or the share of job applications sent to jobs in each 1-digit industry. Columns (4)-(6) correspond to multinomial logit models for the 1-digit occupation of the UI recipients new job. Explanatory variables in these models are dummies for the 1-digit occupation of the previous job or the share of job applications sent to jobs in each 1-digit occupation. Specifications in Columns (1)-(6) exclude individuals who found a job in the smallest industry or occupation, or who had their previous job in this industry or occupation (see footnote \ref{fn:PredValExcl}. Columns (7)-(9) correspond to linear regressions where the outcome variable is the (standardized) firm fixed effect for the UI recipients new job. Explanatory variables in these models are the firm fixed effect for the UI recipients previous job or the average firm fixed effect across all the applied-for jobs. Columns (10)-(12) correspond to linear regressions where the outcome variable is the typical wage of the UI recipients new job. Explanatory variables in these models are the typical wage of the UI recipients previous job or the average typical wage across all the applied-for jobs. The table reports the R^2 for the linear regression models. For the multinomial logit models, the table reports the McFadden's pseudo-R^2. The last row of the table show the p-value for testing the exclusion of all explanatory variables pertaining to applied-for jobs.

Table A.10: Predicting job outcomes from application data vs. prior job characteristics, men only

Job outcome:	Industry (1 digit)			Occupation (1 digit)			Firm wage level			Typical wage		
Model:	Multinomial logit			Multinomial logit			Linear regression			Linear regression		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Explanatory variables												
Characteristics of previous job	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Characteristics of applied-for jobs	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes
# of parameters	99	99	189	63	63	119	2	2	2	2	2	2
(pseudo-)R-squared	0,276	0,249	0,362	0,300	0,402	0,486	0,049	0,094	0,121	0,196	0,267	0,346
p-value, test of excluding applied-for job variables	$< 0{,}01$		$< 0{,}01$	$< 0{,}01$		$< 0{,}01$	$< 0{,}01$		$< 0{,}01$	$< 0{,}01$		$< 0{,}01$

Notes: The table examines the predictiveness of job applications and past job characteristics for the male half of the sample. Columns (1)-(3) correspond to multinomial logit models for the 1-digit industry of the UI recipients new job. Explanatory variables in these models are dummies for the 1-digit industry of the previous job or the share of job applications sent to jobs in each 1-digit industry. Columns (4)-(6) correspond to multinomial logit models for the 1-digit occupation of the UI recipients new job. Explanatory variables in these models are dummies for the 1-digit occupation of the previous job or the share of job applications sent to jobs in each 1-digit occupation. Specifications in Columns (1)-(6) exclude individuals who found a job in the smallest industry or occupation, or who had their previous job in this industry or occupation (see footnote \ref{fn:PredValExcl}. Columns (7)-(9) correspond to linear regressions where the outcome variable is the (standardized) firm fixed effect for the UI recipients new job. Explanatory variables in these models are the firm fixed effect for the UI recipients previous job or the average firm fixed effect across all the applied-for jobs. Columns (10)-(12) correspond to linear regressions where the outcome variable is the typical wage of the UI recipients new job. Explanatory variables in these models are the typical wage of the UI recipients previous job or the average typical wage across all the applied-for jobs. The table reports the R^2 for the linear regression models. For the multinomial logit models, the table reports the McFadden's pseudo-R^2. The last row of the table show the p-value for testing the exclusion of all explanatory variables pertaining to applied-for jobs.

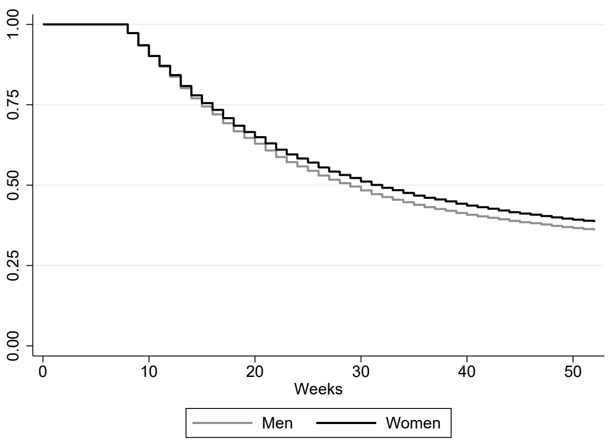
Table A.11: Predicting job outcomes from application data vs. prior job characteristics, women only

Job outcome:	Industry (1 digit)			Occupation (1 digit)			Firm wage level			Typical wage		
Model:	Multinomial logit			Multinomial logit			Linear regression			Linear regression		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Explanatory variables												
Characteristics of previous job	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Characteristics of applied-for jobs	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes
# of parameters	99	99	189	63	63	119	2	2	2	2	2	2
(pseudo-)R-squared	0,271	0,213	0,339	0,328	0,367	0,483	0,063	0,063	0,107	0,307	0,281	0,409
p-value, test of excluding applied-for job variables	$< 0{,}01$		$< 0{,}01$	$< 0{,}01$		$< 0{,}01$	$< 0{,}01$		$< 0{,}01$	$< 0{,}01$		$< 0{,}01$

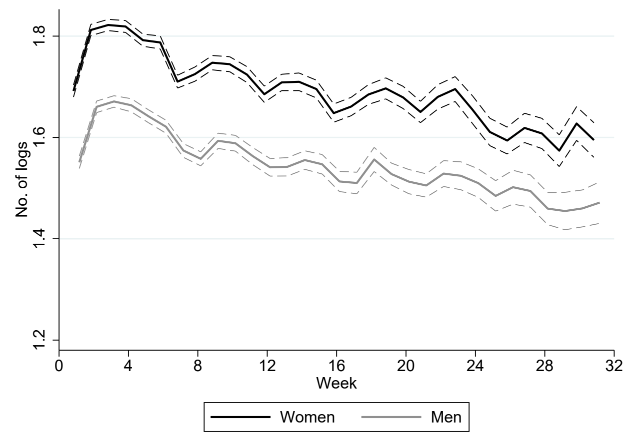
Notes: The table examines the predictiveness of job applications and past job characteristics for the female half of the sample. Columns (1)-(3) correspond to multinomial logit models for the 1-digit industry of the UI recipients new job. Explanatory variables in these models are dummies for the 1-digit industry of the previous job or the share of job applications sent to jobs in each 1-digit industry. Columns (4)-(6) correspond to multinomial logit models for the 1-digit occupation of the UI recipients new job. Explanatory variables in these models are dummies for the 1-digit occupation of the previous job or the share of job applications sent to jobs in each 1-digit occupation. Specifications in Columns (1)-(6) exclude individuals who found a job in the smallest industry or occupation, or who had their previous job in this industry or occupation (see footnote \ref{fn:PredValExcl}. Columns (7)-(9) correspond to linear regressions where the outcome variable is the (standardized) firm fixed effect for the UI recipients new job. Explanatory variables in these models are the firm fixed effect for the UI recipients previous job or the average firm fixed effect across all the applied-for jobs. Columns (10)-(12) correspond to linear regressions where the outcome variable is the typical wage of the UI recipients new job. Explanatory variables in these models are the typical wage of the UI recipients previous job or the average typical wage across all the applied-for jobs. The table reports the R^2 for the linear regression models. For the multinomial logit models, the table reports the McFadden's pseudo-R^2. The last row of the table show the p-value for testing the exclusion of all explanatory variables pertaining to applied-for jobs.

Figure A.1: Survival rates and registered applications

(a) Kaplan-Meier survivor functions in non-employment (unweighted)



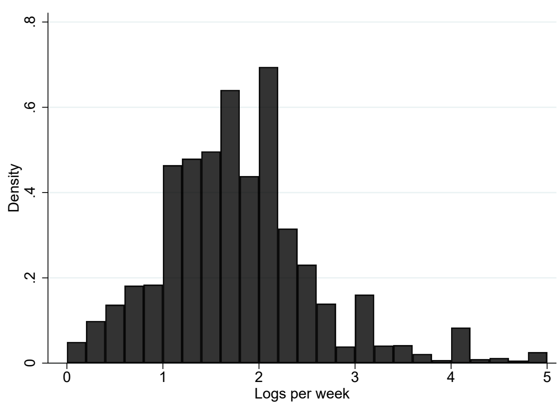
(b) Average number of logged applications



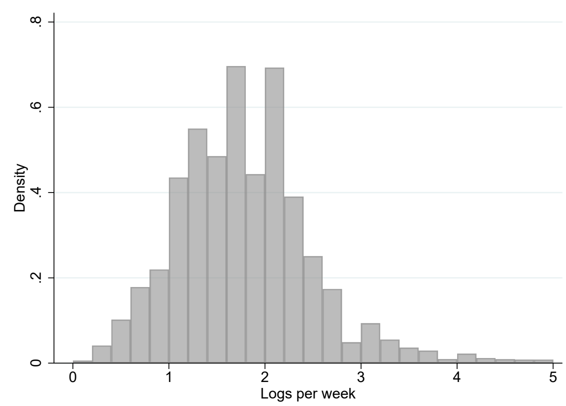
Note: The figures plot gender-specific Kaplan-Meier survival rate in nonemployment estimates (left) and the average number of registered applications (right). The X-axis measure weeks since the start of the UI spell. The Kaplan-Meier estimates of the survivor function in nonemployment are estimated on a version of the main analysis sample where the requirement of finding a job within a year has not been imposed. Average number of registered applications is shown for the main analysis sample.

Figure A.2: Distribution of registered applications per week

(a) Baseline sample

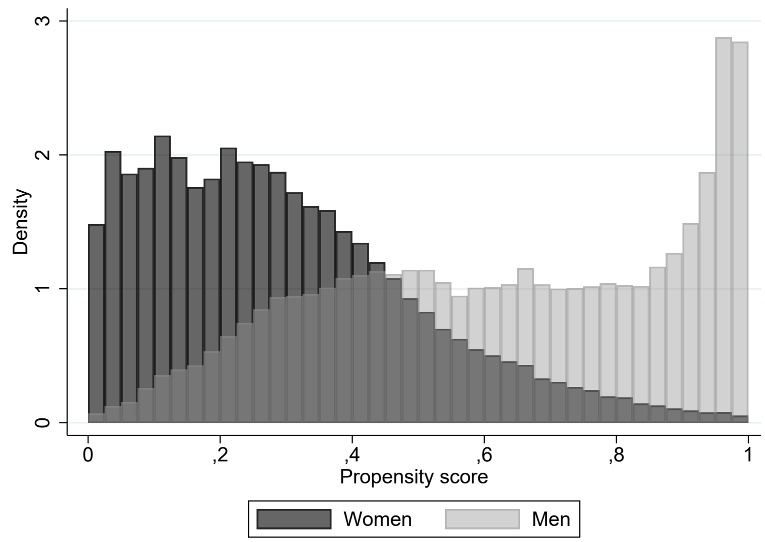


(b) Analysis sample



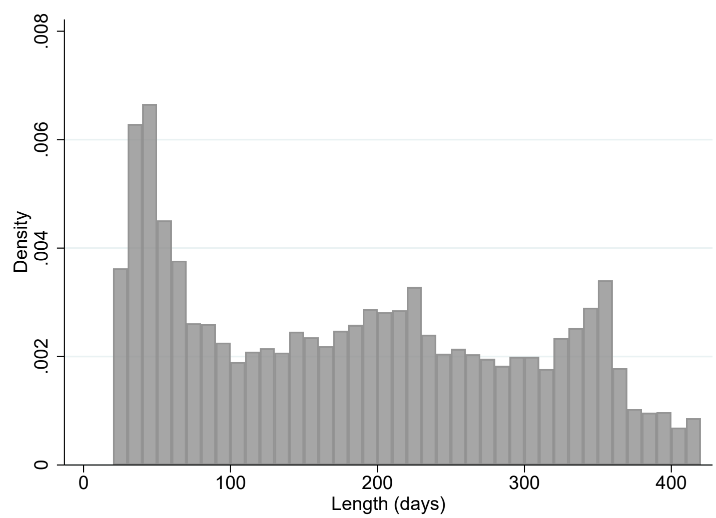
Note: Figure plots the distribution of average applications per week for the unrestricted sample (left) and the analysis sample (right).

Figure A.3: Distribution of estimated propensity scores for descriptive analysis



Note: Figure plots the distribution of male propensity score estimates for men and women. Propensity scores outside the range [0.01,0.99] have been trimmed to avoid extreme weights.

Figure A.4: Parental leave: Distribution of average length



Notes: The figure plots the distribution of the average length of parental leave in days across firms with more than 28 days of leave on average.

B Online Appendix: Robustness checks

B.1 Alternative sample definition

We impose several restrictions on our main sample as laid out in Section 3.1. Below we show that our findings are not sensitive to these restrictions. Figures B.1 to B.5 replicate the female-male gaps in average application and hiring shares for samples based on different sample selection criteria than the ones used in our main analysis. Figure B.1 is based on a sample where we do not exclude the last 4 weeks of applications, whereas in Figure B.2 we focus on unemployment spells that find a job within 26 weeks (in contrast to the 52 week requirement used in the main text). In Figure B.3 we relax our sample restriction of only including unemployment spells with at least 4 registered applications by selecting all spells that have at least one application instead.[64] Figure B.4 removes the restriction of at least 8 weeks of unemployment to enter the sample, and instead include all available unemployment spells in the sample period.[65] Figure B.5 replicates the results for a sample that is not restricted to end in a new hire. Here we treat unemployment as a separate category to the hiring outcomes. Common to all of our robustness tests on the sample selection criteria is that the results do not change qualitatively. In fact, female-male gaps are remarkably stable across the different samples.

B.2 Displacement sample

In this section we repeat our main results for a subsample of individuals who are identified as being a part of a mass-layoff using standard definitions from the literature. The motivation for doing so is to asses the role of differential selection into UI by gender. Our main sample does not have any restrictions on the cause of entry into UI (besides of course UI eligibility) so in principle there may, for example, be slightly more males who enter UI more or less voluntary (through a quit) whereas women to larger extent enter involuntary (layoff). If the cause of entry into UI also affects application and hiring behavior this may contribute to the gender application gap. By focusing on individuals who enter unemployment through a "mass layoff" we try to focus on a group of individuals with a similar cause of entry.

[64] Note that registering at least one application is necessary to appear in the Joblog data we use.

[65] Nevertheless, these unemployment spells need to be at minimum 4 weeks long in order for us to properly identify them, see Section 3.1.

To identify spells in our estimation sample (see Section A.1) which can be linked to a mass layoff we use standard definitions from the literature (see e.g. Lachowska et al. (2020); Bertheau et al. (2022)). We only consider spells where the pre-displacement firm have at least 5 workers employed prior to the displacement event and where they experience a reduction in firm size of more than 30 percent between the displacement year (i.e. the year where the unemployment spell start) and the subsequent year.[66] We focus on all spells who do not return to the pre-displacement firm within our sample window, and we further require that the worker worked in the pre-displacement firm for at least 2 consecutive years prior to unemployment. In the end we identify 6217 spells (around 6 percent of our sample after trimming) satisfying the above criteria and for these spells we then proceed by creating application and hiring gaps.

In figure B.6 we report application and hiring gaps for the subsample of spells which can be linked to a mass layoff. Compared to Figure 1 in the main text we see broadly similar patterns , however precision is markedly lower in the displacement sample due to the reduction in sample size. This suggests that our main results are not explained by differential selection at entry into UI.

B.3 Conditioning on different observables

In the main text, our descriptive analysis conditions on observables by propensity score reweighting on a set of variables selected through a LASSO procedure. We have, however, also experimented with several other approaches to conditining on observables. None of these change our conclusions. In this section we present results from some of these alternative approaches.

In Figure B.7 we present results after propensity score reweighting only on the 3-digit industry of the previous job, thus imposing exact balance on previous industry across men and women. Similarly, in Figure B.8 we present results after propensity score reweighting only on the 3-digit occupation of the previous job, thus imposing exact balance on previous occupation. Finally, in Figure B.9 in which we use the same set of conditioning variables as the main analysis but include dummies for the quarter of entry into UI in the propensity score estimation, see Section 3.4. The purpose of this is to control for seasonality, i.e. whether entering the sample at different times is

[66] Since average firm size is low in Denmark, requiring firms to have more employees would reduce the sample size quite dramatically.

important for the differences in application behavior and hiring outcomes we observe.

Throughout the various alternative approaches, we see a similar pattern of gender gaps as in the main text.

B.4 Raw gender gaps in application and hiring outcomes

In Figure B.10, we show raw gender gaps in applications and hiring outcomes without conditioning on unobservables. We see that the overall patterns are similar to the conditional results presented in the main text but that, unsurprisingly, the raw gaps tend to be much larger in magnitude.

B.5 Gender gaps in number of applications instead of shares

In our main analysis we measure gender gaps in the share of applications going to different jobs. In Figure B.11 we instead show gender gaps in the absolute number of applications sent to different jobs. We see that the overall patterns of results from the main analysis remains.

B.6 Gender-specific measures of wages

B.6.1 Jobs' typical wage level

One drawback of the measure of a job's typical wage level that we use in the main analysis is that it does not allow for the possibility that men and women face different wages in the same job. A particular concern here is the possibility that some types of jobs tend to pay high wages to women but not men, while for other jobs it is the reverse. If this is the case we might expect women to apply much more for jobs in which they face higher wages and vice versa for men.

To address this possibility, this section presents results based on two alternative measures wages that capture the typical wage level faced by either women or men in a given job. We construct these measures simply by repeating the last step of the wage prediction exercise underlying our typical wage only for the male or female half of the sample (see Appendix A.4).

Figure B.12 shows gender gaps in applications and hiring splitting jobs according to the typical wage level they pay to either women or men. Although the exact size of the application and hiring gaps change slightly, the overall pattern of results is identical to what we see in the main analysis.

B.6.2 Gender-specific AKM firm effects

In similar spirit to Section B.6 this section contains results based on gender specific measures of firm effects thereby analyzing the importance of potential gender differences in firm effects and thus rankings across firms. We construct these measures simply by repeating our AKM estimation only for the male or female half of our AKM estimation sample (see Appendix A.3). After estimation we then rank all new jobs and applications according to the estimated firm effects and construct application and hiring gaps. The results are presented in Figure B.13. We see very similar overall patterns regardless of whether we rank based on firm effects estimated for men or women only. Although the exact size of the application and hiring gaps change slightly, the overall pattern of results is very similar to what we see in the main analysis. Note that precision is lower since the set of connected firms (i.e. firms for whom we can actually estimate firm effects) is substantially reduced when we restrict the sample to only males or females.[67]

B.7 Decomposing raw gender gaps

In this section, we present a decomposition of the raw gender gaps in hiring outcomes. We do this by simply forgoing the first step of the decomposition used in the main analysis: starting from the raw gender gaps in hiring outcomes in the data, we propensity score reweight the women in the sample to the same application behavior as men to see how much of this raw gap applications can explain.

Table B.1 shows the decomposition. Unsurprisingly, we see that raw gender gaps before conditioning on observables are generally larger than the baseline hiring gaps considered in the main analysis. We also see that applications are capable of explaining a larger share of these raw gaps. The only exception to this is firm-wage level, where the raw gap is smaller than the baseline gap and where applications are capable of explaining a smaller share of the raw gap.

[67] Coverage is lower for firm effects estimated for women only. Compared to the firm effects estimated based on the full sample we lose 33 percent of firm effects when estimation is based on women only. The same number for men is 14 percent.

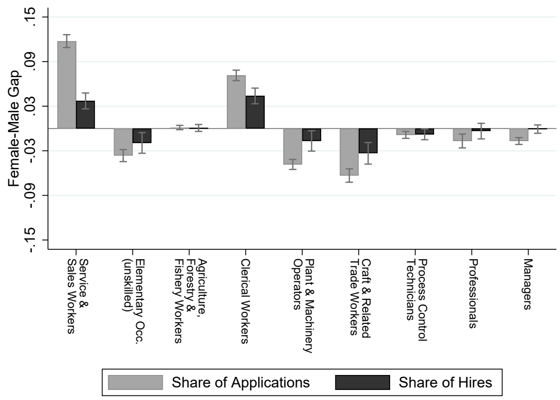
Table B.1: Decomposing gender gaps without conditioning on observables

	Baseline	**Explained by application**	**Residual**
Occupational segregation	0.228	0.178	0.050
(Duncan index, 1-digit)	*(0.003)*	*(0.006)*	*(0.006)*
		[0.78]	[0.22]
Occupational segregation	0.351	0.225	0.126
(Duncan index, 2-digit)	*(0.003)*	*(0.007)*	*(0.006)*
		[0.64]	[0.36]
Industry segregation	0.253	0.218	0.036
(Duncan index, 1-digit)	*(0.003)*	*(0.006)*	*(0.006)*
		[0.86]	[0.14]
Industry segregation	0.263	0.205	0.058
(Duncan index, 2-digit)	*(0.003)*	*(0.006)*	*(0.006)*
		[0.78]	[0.22]
Firm wage level	0.026	0.012	0.014
(Male-female gap, std. AKM fixed effect)	*(0.004)*	*(0.017)*	*(0.017)*
		[0.46]	[0.54]
Firm wage level	0.015	0.012	0.003
(Male-female gap, AKM fixed effect)	*(0.001)*	*(0.002)*	*(0.001)*
		[0.81]	[0.19]
Typical wage for job	0.039	0.034	0.005
(Male-female gap, log typical wage)	*(0.001)*	*(0.002)*	*(0.002)*
		[0.87]	[0.13]
Actual Wages	0.069	0.055	0.014
(Male-female gap, log wage)	*(0.002)*	*(0.003)*	*(0.003)*
		[0.80]	[0.20]

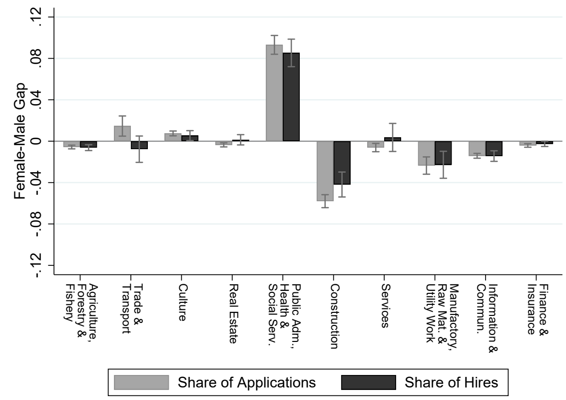
Notes: The table decomposes the raw gaps in hiring outcomes. The gaps are decomposed into a part explained by applications and a residual gap. Standard errors based on bootstrapping individuals are shown in parenthesis. Brackets report the share of the raw gap explained by each component. To ease comparability, the sample is the same as the sample used for the decomposition in Table 4.

Figure B.1: Gender gaps in applications and hiring outcomes, no exclusion of last 4 weeks

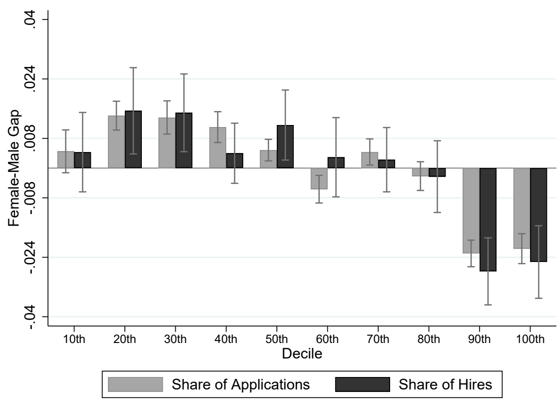
(a) Occupations



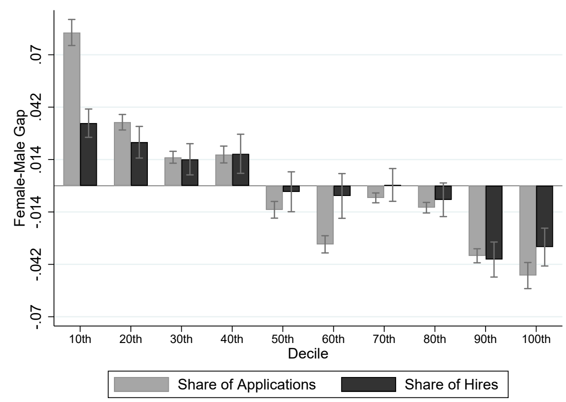
(b) Industries



(c) Firm wage level

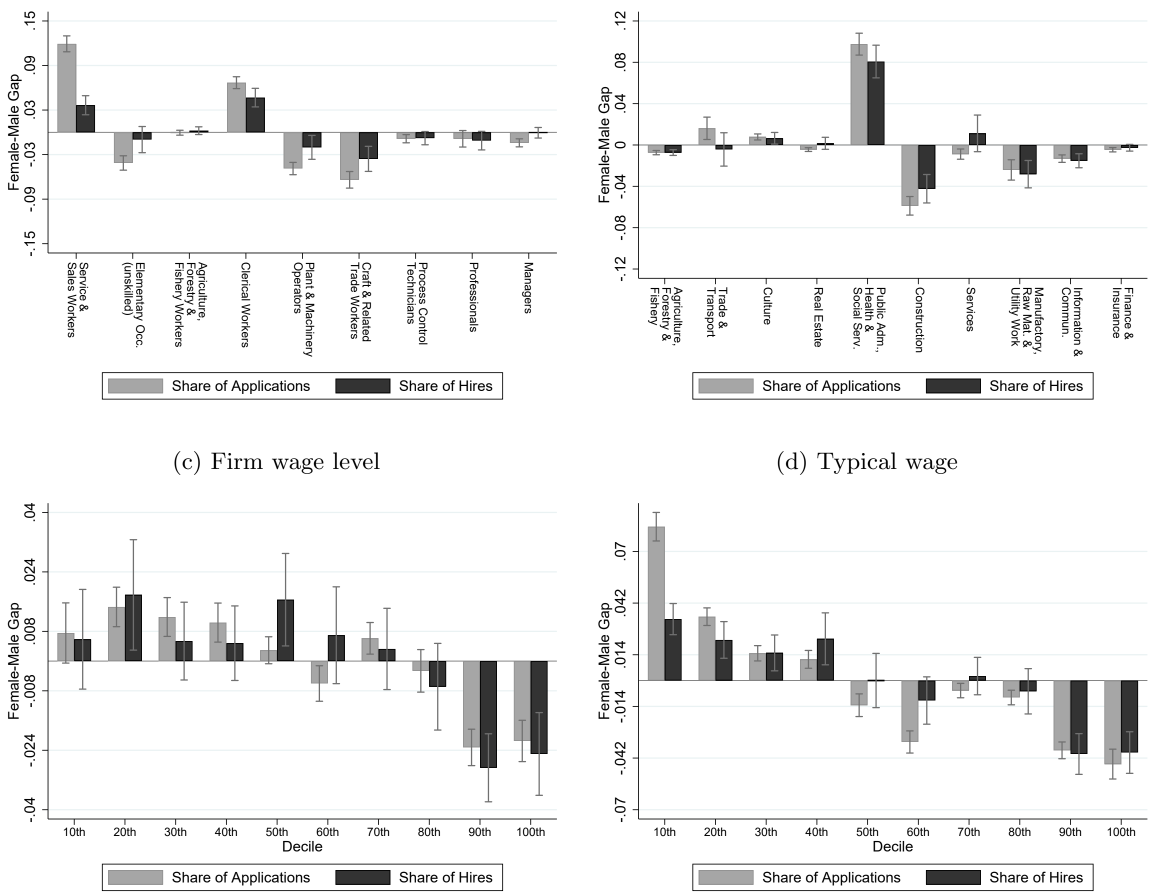


(d) Typical wage



Notes: The figures plot gender gaps in the share of applications going to specific types of jobs and corresponding gender gaps in hiring shares. The figure relaxes the sample restriction on the last 4 weeks of unemployment and consider all applications. All gaps are based on a reweighted sample so are conditional on observables. Occupations and industries on the x-axis are ordered from lowest paying on the left to highest paying on the right. The 95 percent confidence bars are based on standard errors clustered on the individual level.

Figure B.2: Gender gaps in applications and hiring outcomes, 26 week sample



(a) Occupations

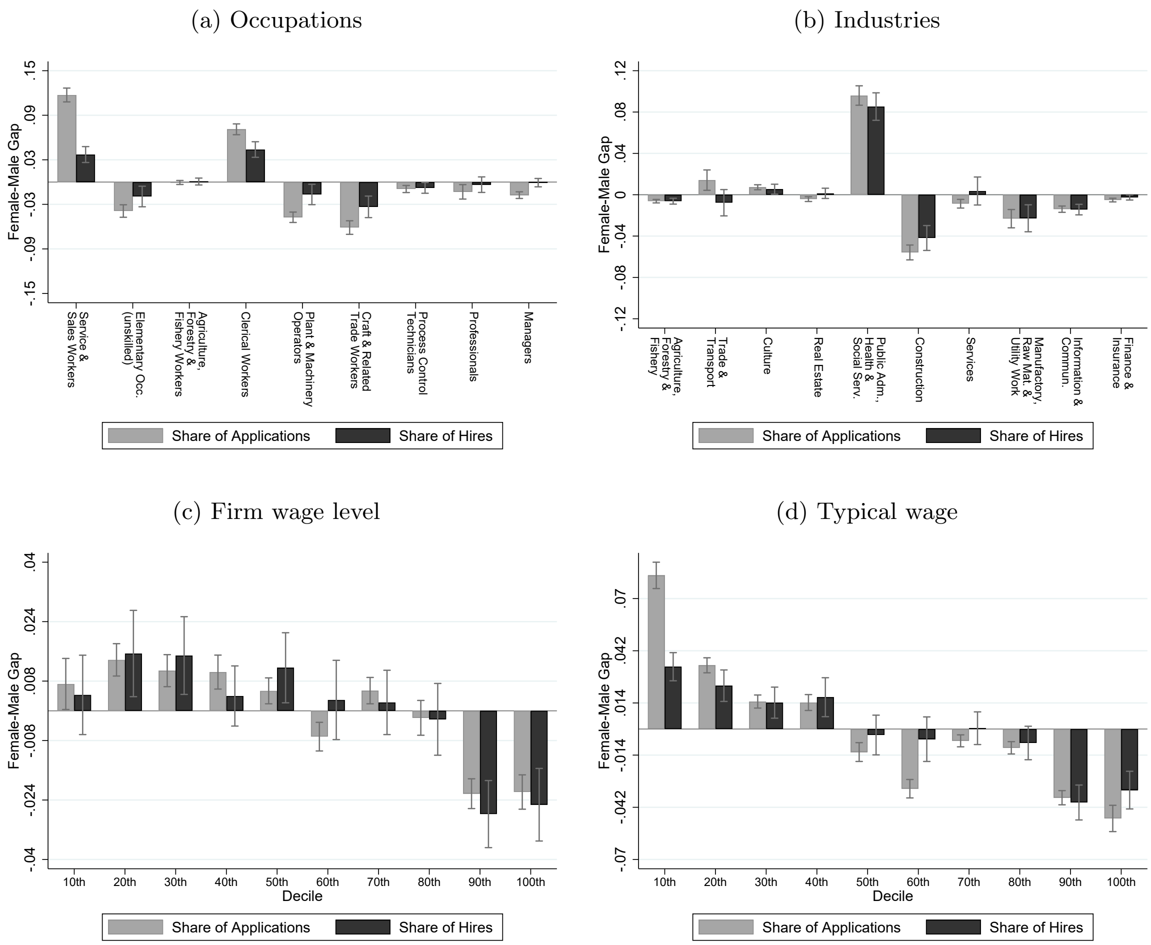
(b) Industries

(c) Firm wage level

(d) Typical wage

Notes: The figures plot gender gaps in the share of applications going to specific types of jobs and corresponding gender gaps in hiring shares. The figure changes the sample to only consider spells lasting at most 26 weeks. All gaps are based on a reweighted sample so are conditional on observables. Occupations and industries on the x-axis are ordered from lowest paying on the left to highest paying on the right. The 95 percent confidence bars are based on standard errors clustered on the individual level.

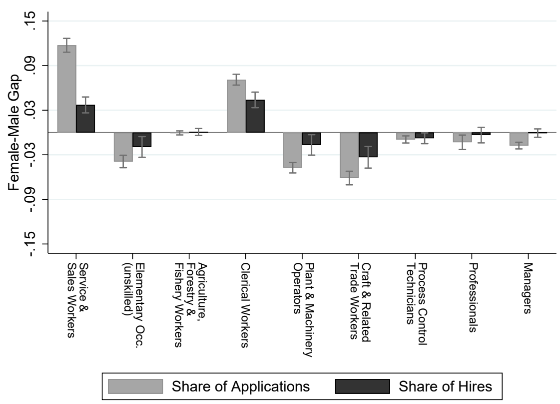
Figure B.3: Gender gaps in applications and hiring outcomes, only 1 application registration requirement



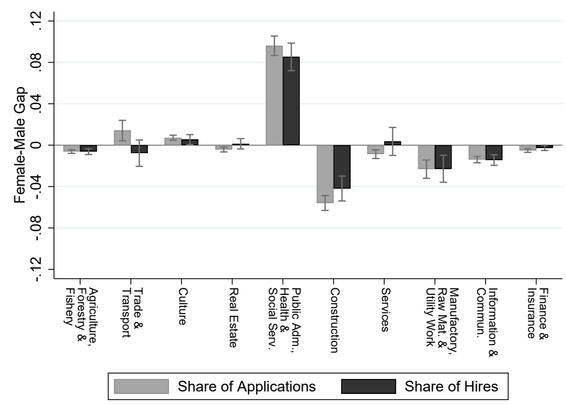
Notes: The figures plot gender gaps in the share of applications going to specific types of jobs and corresponding gender gaps in hiring shares. The figure relaxes the samples requirement of having min. 4 registered applications to at least 1 application. All gaps are based on a reweighted sample so are conditional on observables. Occupations and industries on the x-axis are ordered from lowest paying on the left to highest paying on the right. The 95 percent confidence bars are based on standard errors clustered on the individual level.

Figure B.4: Gender gaps in applications and hiring outcomes, no spell length requirement

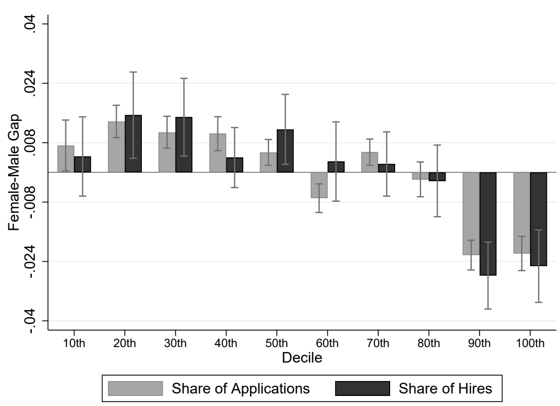
(a) Occupations



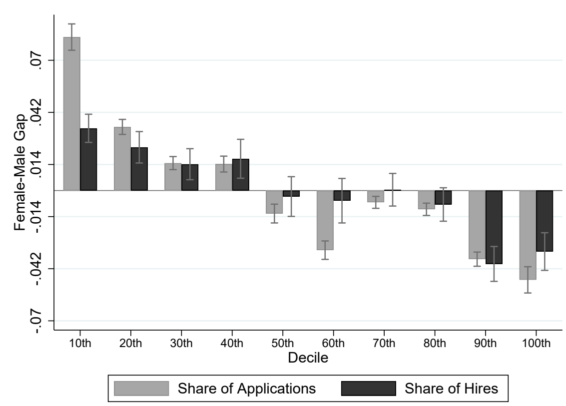
(b) Industries



(c) Firm wage level



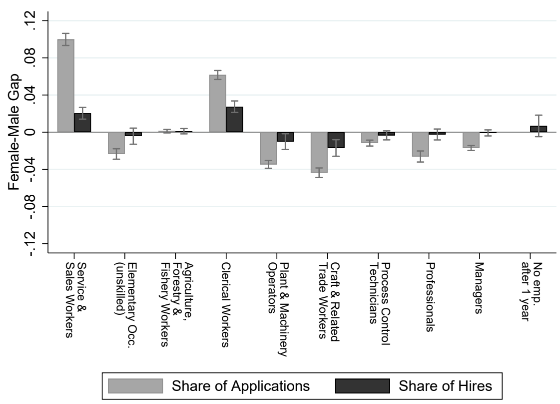
(d) Typical wage



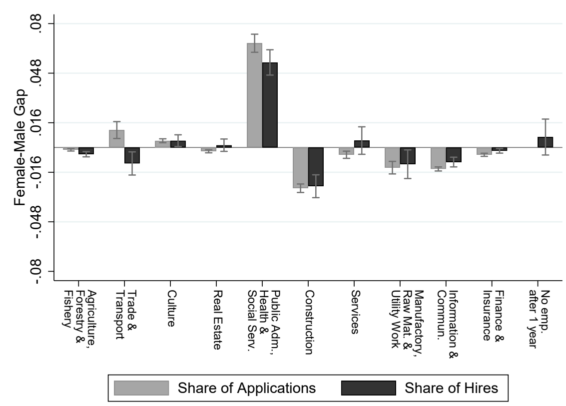
Notes: The figures plot gender gaps in the share of applications going to specific types of jobs and corresponding gender gaps in hiring shares. The figure relaxes the sample restriction on min. 8 weeks spell length and considers all spells. All gaps are based on a reweighted sample so are conditional on observables. Occupations and industries on the x-axis are ordered from lowest paying on the left to highest paying on the right. The 95 percent confidence bars are based on standard errors clustered on the individual level.

Figure B.5: Gender gaps in applications and hiring outcomes, no employment requirement

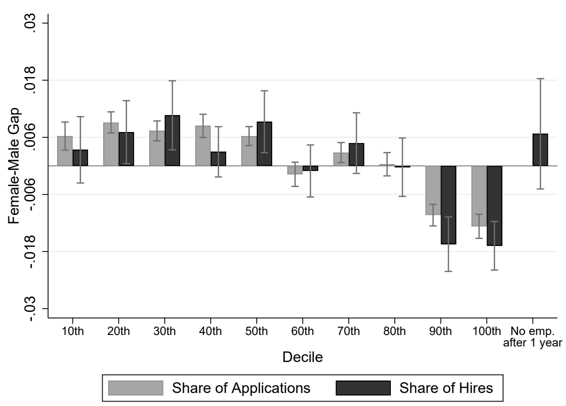
(a) Occupations



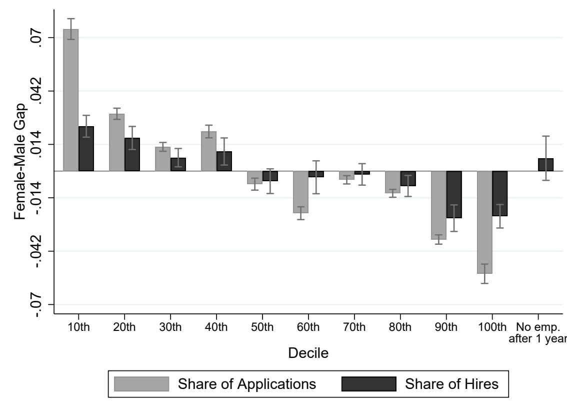
(b) Industries



(c) Firm wage level

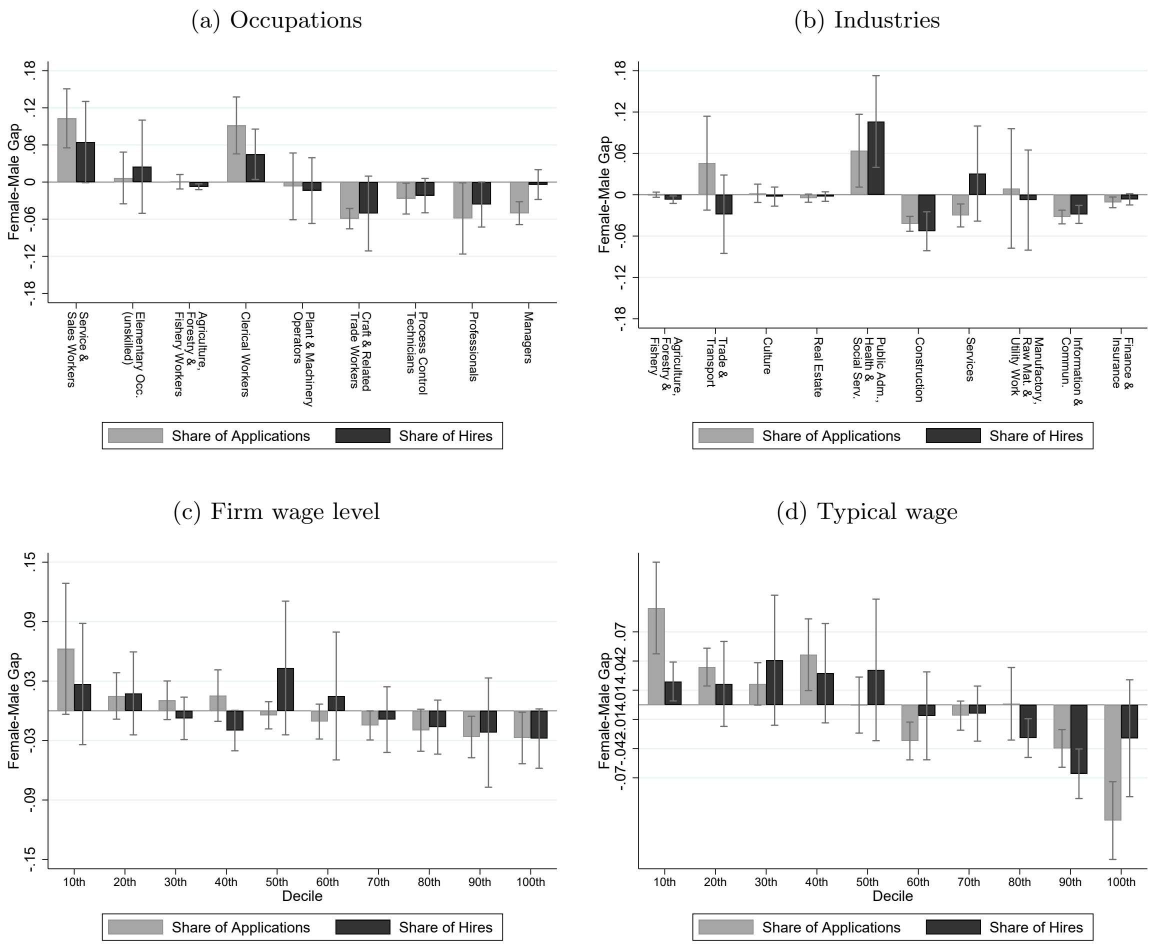


(d) Typical wage



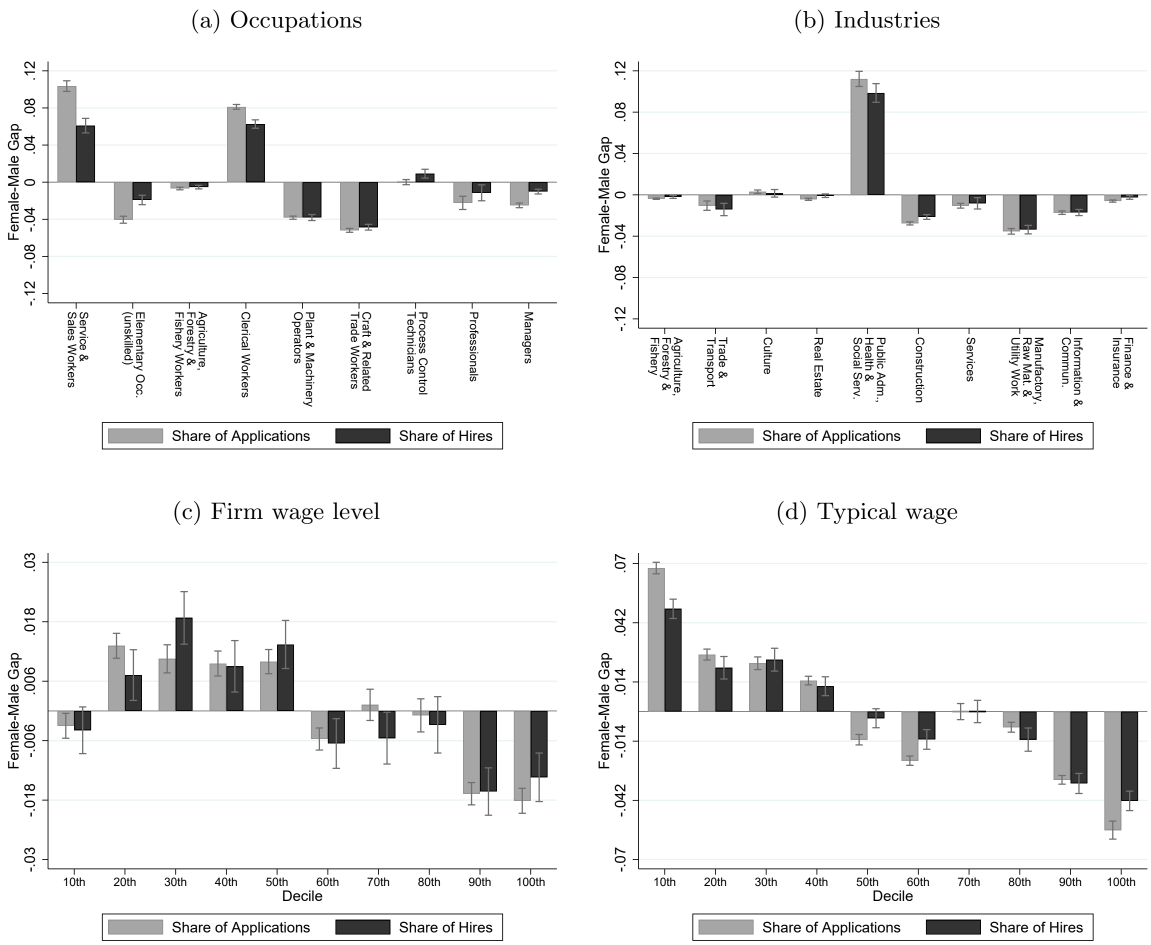
Notes: The figures plot gender gaps in the share of applications going to specific types of jobs and corresponding gender gaps in hiring shares. The sample also includes spells that do not end in employment. Staying unemployed is treated as a separate category throughout. All gaps are based on a reweighted sample so are conditional on observables. Occupations and industries on the x-axis are ordered from lowest paying on the left to highest paying on the right. The 95 percent confidence bars are based on standard errors clustered on the individual level.

Figure B.6: Gender gaps in applications and hiring outcomes, conditioning only on previous industry



Notes: The figures plot gender gaps in the share of applications going to specific types of jobs and corresponding gender gaps in hiring shares. The sample consists of spells who can be linked to a mass-layoff, see Section B.2. All gaps are computed after propensity score reweigthing on only the 3-digit industry of the previous job. Occupations and industries on the x-axis are ordered from lowest paying on the left to highest paying on the right. The 95 percent confidence bars are based on standard errors clustered on the individual level.

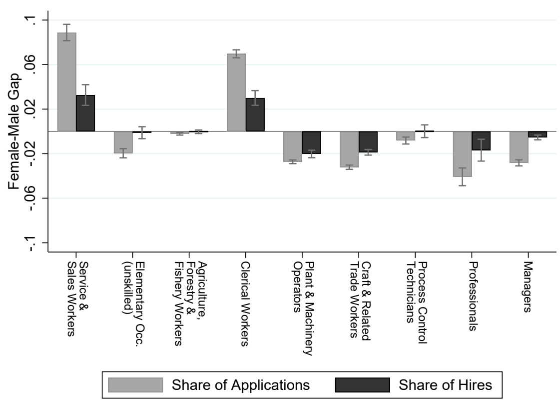
Figure B.7: Gender gaps in applications and hiring outcomes, conditioning only on previous industry



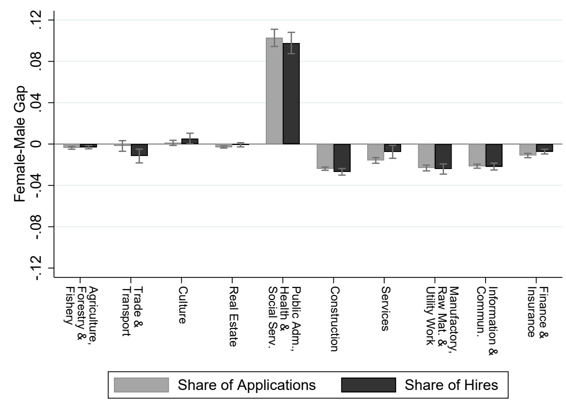
Notes: The figures plot gender gaps in the share of applications going to specific types of jobs and corresponding gender gaps in hiring shares. All gaps are computed after propensity score reweigthing on only the 3-digit industry of the previous job. Occupations and industries on the x-axis are ordered from lowest paying on the left to highest paying on the right. The 95 percent confidence bars are based on standard errors clustered on the individual level.

Figure B.8: Gender gaps in applications and hiring outcomes, conditioning only on previous occupation

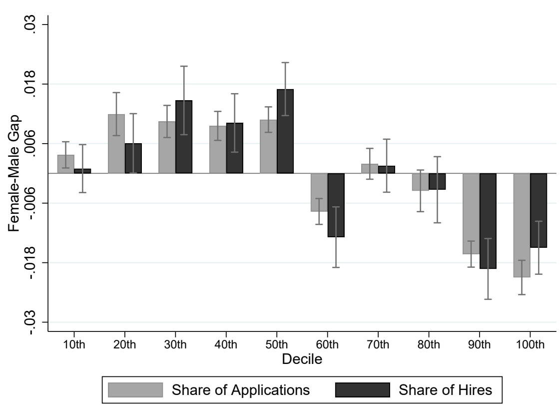
(a) Occupations



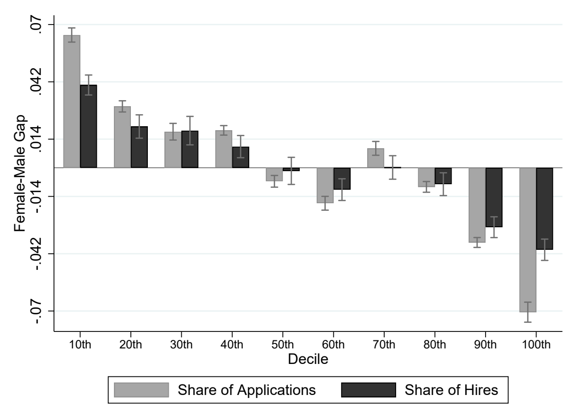
(b) Industries



(c) Firm wage level



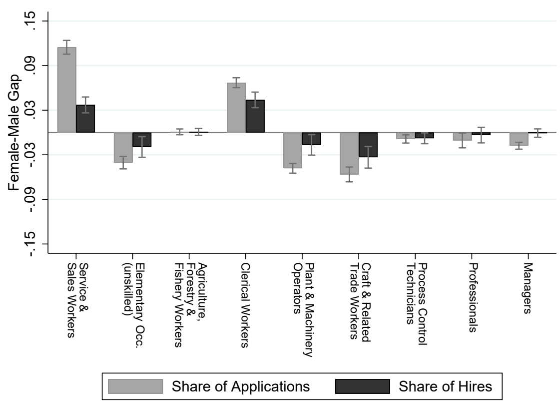
(d) Typical wage



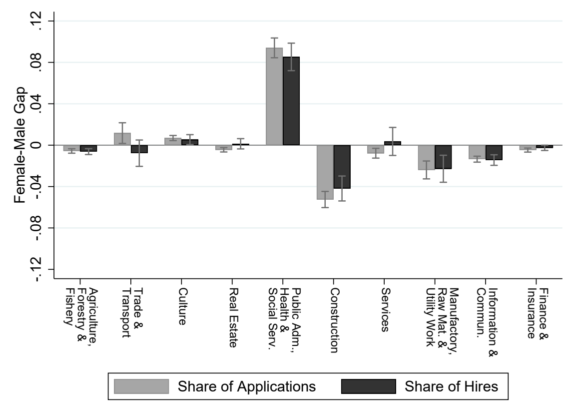
Notes: The figures plot gender gaps in the share of applications going to specific types of jobs and corresponding gender gaps in hiring shares. All gaps are computed after propensity score reweigthing on only the 3-digit occupation industry of the previous job. Occupations and industries on the x-axis are ordered from lowest paying on the left to highest paying on the right. The 95 percent confidence bars are based on standard errors clustered on the individual level.

Figure B.9: Gender gaps in applications and hiring outcomes, seasonality controls

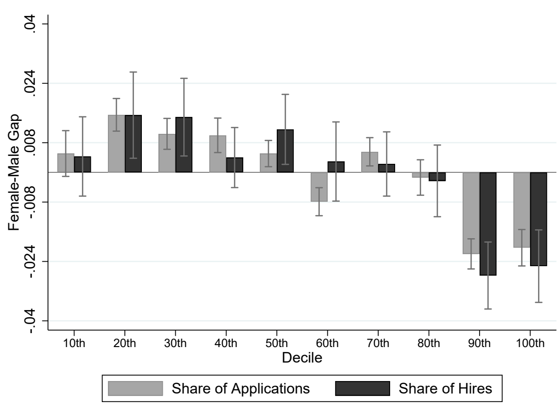
(a) Occupations



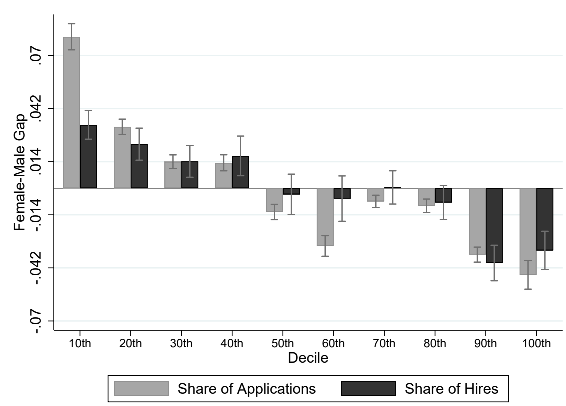
(b) Industries



(c) Firm wage level



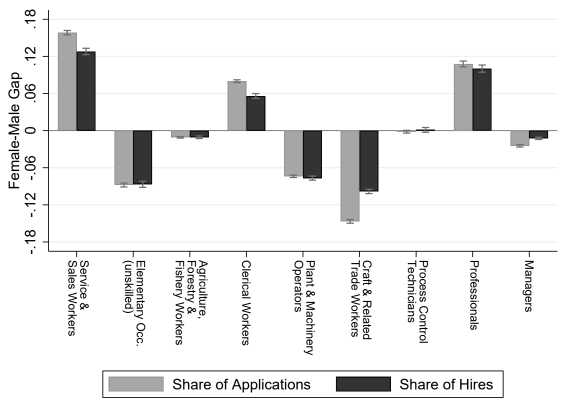
(d) Typical wage



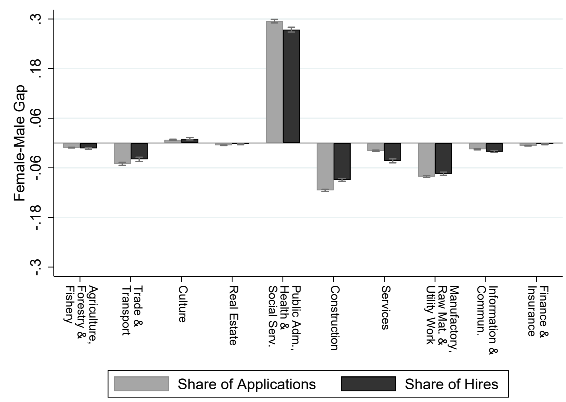
Notes: The figures plot gender gaps in the share of applications going to specific types of jobs and corresponding gender gaps in hiring shares. All gaps are computed after propensity score reweigthing on the quarter of inflow into unemployment along with all the conditioning variables used in the main analysis. Occupations and industries on the x-axis are ordered from lowest paying on the left to highest paying on the right. The percent confidence bars are based on standard errors clustered on the individual level.

Figure B.10: Gender gaps in applications and hiring outcomes, raw

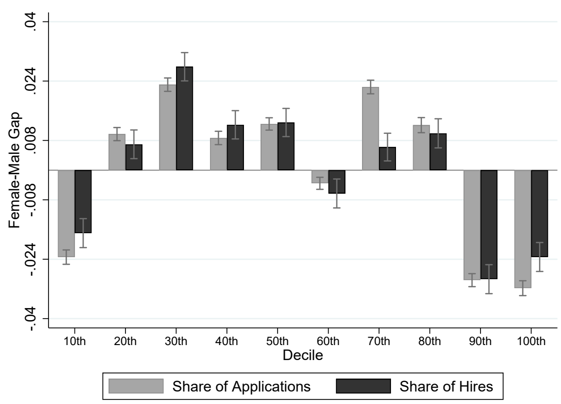
(a) Occupations



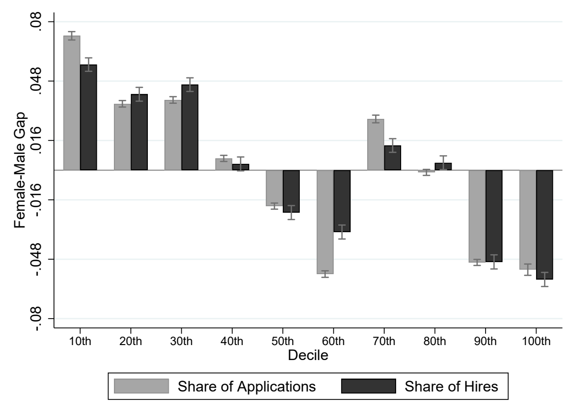
(b) Industries



(c) Firm wage level

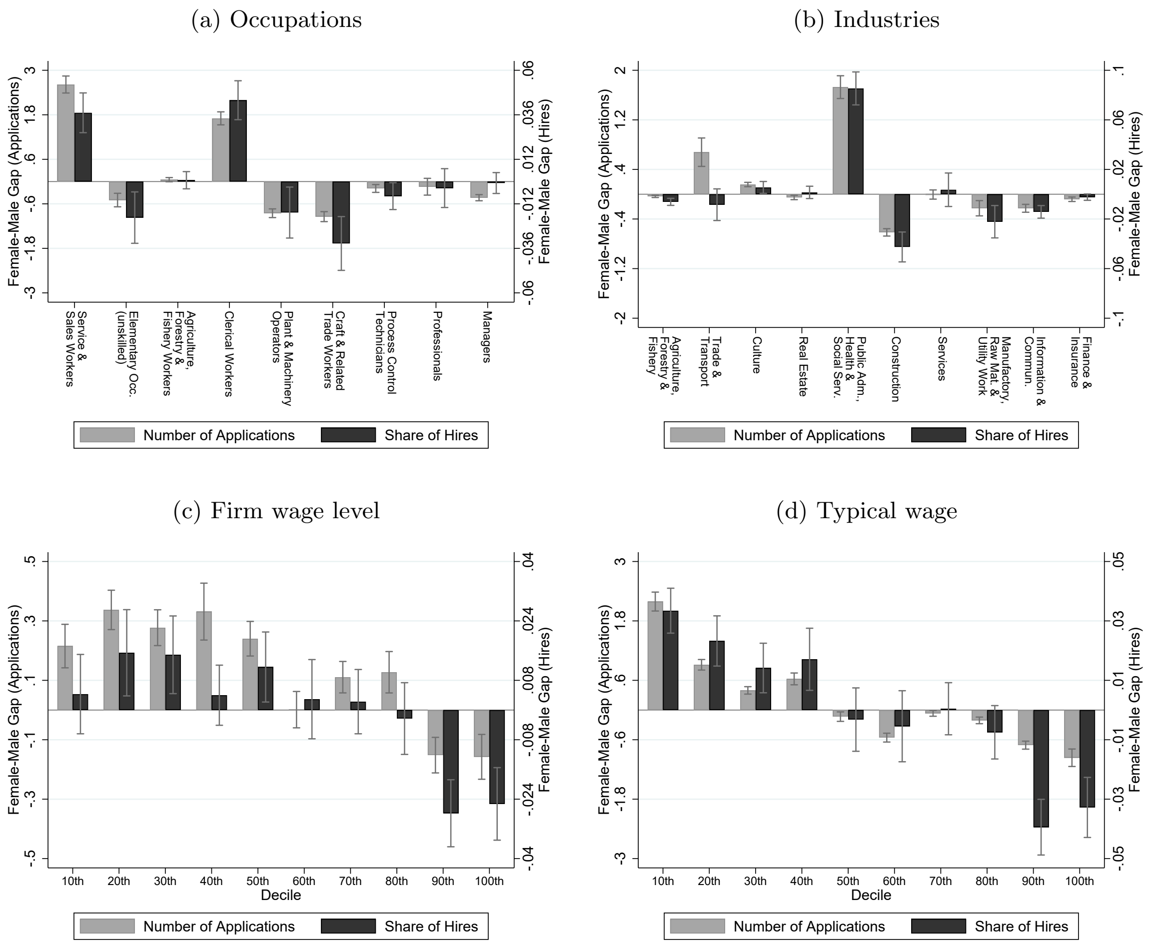


(d) Typical wage



Notes: The figures plot gender gaps in the share of applications going to different types of jobs along with corresponding gender gaps in hiring shares. The figure shows raw gaps without conditioning on observables. Occupations and industries on the x-axis are ordered from lowest paying on the left to highest paying on the right. The 95 percent confidence bars are based on standard errors clustered on the individual level.

Figure B.11: Gender gaps in applications and hiring outcomes, absolute measure for applications

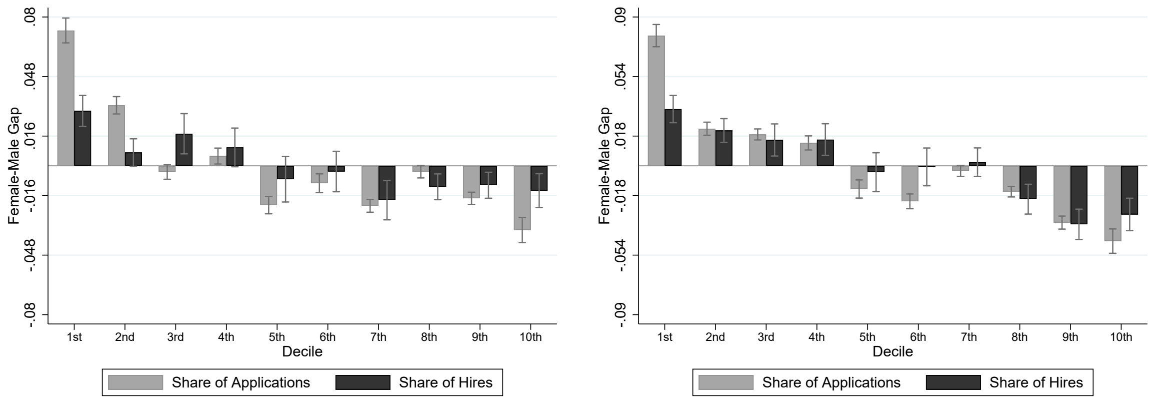


Notes: The figures plot gender gaps in the number of applications going to different types of jobs (left axis) along with corresponding gender gaps in hiring shares (right axis). All gaps are computed after propensity score reweigthing so are conditional on observables. Occupations and industries on the x-axis are ordered from lowest paying on the left to highest paying on the right. The 95 percent confidence bars are based on standard errors clustered on the individual level.

Figure B.12: Gender gaps in job applications and hiring, gender-specific typical wage measures

(a) Typical wage paid women

(b) Typical wage paid to men

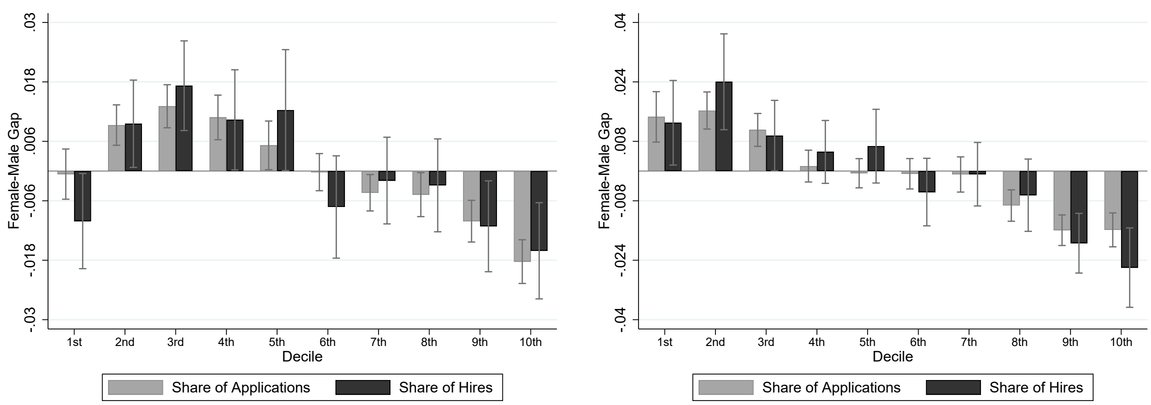


Notes: The figures plot gender gaps in the share of applications going different jobs with different level of typical wages paid to women (left) and men (right). All gaps are computed after propensity score reweigthing so are conditional on observables. Occupations and industries on the x-axis are ordered from lowest paying on the left to highest paying on the right. The 95 percent confidence bars are based on standard errors clustered on the individual level.

Figure B.13: Gender gaps in job applications and hiring, gender-specific firm effects

(a) Firm effects based on women only

(b) Firm effects based on men only



Notes: The figures plot gender gaps in the share of applications going different jobs with different level of firm effects estimated only on a sample of women (left) or men (right). Subsequently the obtained firm effects are then used to rank all applications and jobs. All gaps are computed after propensity score reweigthing so are conditional on observables. See Section A.3 regarding the estimation sample for the firm effects. The 95 percent confidence bars are based on standard errors clustered on the individual level.

C Online Appendix: Additional results

C.1 Changes in application behavior over time

In our analysis we pool all applications sent during the unemployment spell and analyze the composition of this pool across different individuals. In principle, however, there could be interesting gender differences in how application behavior changes over the course of an unemployment spell. To present some evidence on the importance of such dynamics we create a panel date that for each person and unemployment spell contains an observation that for each month of the spell. We then run an twoway fixed effect regression that uses the monthly average typical wage of applied-for jobs as the outcome variable and includes dummies for the number of months since entering unemployment alongside person-by-spell fixed effects.[68] We run this separately for men and women. The results of this regression shows how men and women change their application behavior over time within an unemployment spell.

Figure C.1 shows the estimates of gender specific time profile of job search from this regression. Changes in application behavior over time are modest and are very similar for men and women. Examining gender gaps in applications at different times throughout the unemployment spell therefore yields very similar results to the pooled results presented in the main text. This is true also for other job characteristcs besides wages.

C.2 Occupation and wages of applied-for jobs relative to previous job

In Figure C.2 we show gender application and hiring gaps when grouping applications and hires by whether the typical wage is higher or lower than the previous job. The figure reveal a stark difference between women and men, where men target higher typical wages compared to the typical wage of their previous job.

In Figure C.3 we repeat the same exercise for occupations. Specifically, we rank occupations according to their average wage and then compare applied-for occupations' rank to the rank of the previous occupation. We again see stark gender differences in the extent to which men and women apply "up" or "down" relative to their previous job. In particular women are more likely to submit

[68] We use the month of entry into unemployment as the baseline and omit a dummy for this month in the regression. We let the fixed effect be specific to each unemployment spell because some individuals in our data show up with more than one unemployment spell. We cluster standard errors at the level of the unemployment spell.

applications which would involve occupational downgrading and are also more likely to be hired in this type of jobs.

C.3 Applications to public vs. private sector jobs

In Figure C.4 we report the application and hiring gaps when we group firms by whether they are private or public.[69] Consistent with previous work on gender and public sector jobs, we see that women send about 9 percentage points more of their applications to public sector jobs and are also about 9 percentage points more likely to be hired into these jobs.

C.4 Decomposing hiring gaps across deciles of the firm and wage distribution

Mirroring the decomposition results for individual industries and occupations in Section 5.4, Figure C.5 decomposes hiring gaps into each decile of the distribution of firm wage-levels (Panel A) and typical wages (Panel B). For almost all deciles, gender differences in applications are able to explain the majority of the baseline hiring gap.

C.5 Comparison to audit study estimates of occupation-specific gender discrimination

For understanding our decomposition results across occupations, varying degrees of direct gender discrimination is a particularly salient possible explanation. A natural way of examining this explanation is to compare our residual hiring gaps to direct estimates of gender discrimination from the audit study literature (Riach and Rich (2002); Rich (2014); Neumark (2018)). Audit studies are controlled experiments that send fictitious job applications to actual vacancies to measure gender discrimination between otherwise identical applicants.

Unfortunately, no existing audit study (or set of studies) provide reliable estimates of discrimination across industries or occupations in Denmark.[70] The audit study data collected by Ahmed et al. (2021), however, provides a useful benchmark. Ahmed et al. (2021) combine data from three audit studies conducted in Sweden between 2016 and 2019 thus the setting and time period is

[69] We distinguish firms by whether they are private or public through their industry affiliation and classify public administration, education, health, culture and services as public firms.

[70] The closest is the Danish audit study of Dahl and Krog (2018). Since they only sent applications to 400 different jobs, however, their sample size is not well suited to construct occupation-specific measures of discrimination.

comparable to our data. The data contain 3,214 applications, covering 15 different types of jobs. Because the gender of the applicant is randomly assigned, gender discrimination in the initial hiring stage can be measured by looking at how often female and male applications received a positive reply. We refer to as the gender gap in callback rates (although note that employer responses were not restricted to come in the form of phone calls).

To construct occupation-level measures of the gender gap in callback rates, we need to map the 15 jobs considered in the audit studies to the occupation groups used in our analysis. Table C.1 shows the result of this mapping. For 12 of the 15 jobs considered, we are uniquely able to map them to 5 different of the occupation group in our data. For the remaining 3 jobs, it is not possible to uniquely map them to an occupation group.

Figure 4a in the main text plots the resulting occupation-specific gender gaps in callback rates against the occupation-level residual hiring gaps from our decomposition.

C.6 Do women apply for jobs with lower wage growth?

In Section 3.5, we saw that men and women apply to jobs with different typical wages. These differences largely correspond to differences in the eventual hiring outcome, leading to the observed gender wage gaps in typical wages. As described in Section, 3.2, however, the wage measures used in our main analysis reflect the wage level at the start of a workers new job. If there are also substantial differences in how fast wages increase over the initial years in a firm these gender gaps may further change over time during employment spells.

To get a sense on whether this is important, this section constructs a measure of the wage growth that a given job offers and examine whether there are gender gaps in applications and hiring also into high vs low wage growth jobs. Specifically, we calculate firm specific 1 and 5 year wage growth rates based on all jobs starting between 2008 and 2016 in a given firm. In order to separate general time trends from the growth rates as well as structural differences between industries, we control for year by industry fixed effects. As many individuals will have left their jobs by the one and, especially, the five year mark, we censor individuals that are no longer employed by the firm at these times.[71]

[71] Obviously it is not random who stays in a firm up to 5 years, and our numbers may therefore partly be driven by dynamic selection. Results should be interpreted with this in mind.

In Figure C.6 we split firms into deciles based on the typical wage growth they offer over either 1 or 5 years. We see that women are more likely to apply to firms with lower wage growth rates, with the exception of the lowest decile. Likewise, men apply substantially more to those firms in the very top deciles. In addition to the fact that women are applying and getting hired more into jobs with lower staring wages, women thus are thus also applying and getting hired into firms that offer lower rates of wage growth. These differences in applications and hiring outcomes contribute to a widening of the gender gap over time.

C.7 Are gender gaps in applications related to motherhood?

Several recent papers have emphasized that gender gaps in the labor market are particularly related to motherhood and its effects on the valuation of non-wage job characteristics. In Figures C.7 and C.8, we attempt shed some light on how motherhood relates to gender differences in job applications. In these figures, we limit our sample to UI recipients in an age window around the prime childbearing years, specifically 25-40 years. We then repeat our descriptive analysis of gender gaps in application and hiring separately for men and women with young children (0-5 years) and for men and women without children. For most of the non-wage job characteristics considered in Section 6.1 and for typical wages, we see that gender gaps in both applications and hiring outcomes tend to be larger when comparing men and women with young children. At the same time, however, we note that very substantial gender differences do exist in both groups.

C.8 Gender differences in the returns to applications (self-fulfilling discrimination)

One explanation for the observed gender application gap is a version of the "self-fulfilling discrimination" mechanism that has been proposed and documented in other settings (Lundberg and Startz, 1983; Coate and Loury, 1993; Glover et al., 2017). In the context of the job application process, the basic idea behind this is as follows: Gender discrimination in hiring implies that there are gender differences in the likelihood that an application turns into a hire for some jobs. As a result women have an incentive to apply less to these jobs and more to jobs where the chance of being hired is higher. In this way, gender differences in the likelihood than an application turns into a hire may explain why there are gender differences in applications.

To check whether we can detect this phenomenon in our data, we first construct measures of gender differences in the likelihood that an application turns into a hire. We do this in the context of a simple linear regression. For some individual in our analysis sample, let y denote a type of job (an occupation, industry or decile of a wage distribution), let a^y be the share of their applications that the individual sent to jobs of type y, and let d^y be an indicator for whether the individual was hired into a job of type y. We then consider estimating the following regression on our weighted analysis sample:

$$d^y = \beta_0^y + \beta_1^y a^y + \varepsilon \tag{10}$$

If a^y is measured in percentage points, the coefficient β_1^y in this regression captures how much the likelihood of being hired into job type y increases if one additional percentage point of applications is targeted to this type of job. If given a causal interpretation, this is a measure of the returns to applications for job type y. To obtain a simple measure of gender differences in the likelihood that an application turns into a hire for job type y, we therefore estimate equation 10 separately for men and women and compute the difference in the estimate of β_1^y across genders. We refer to this as the *gender gap in returns to applications* for job type y. We note that there are obvious concerns with treating the estimate of β_1^y as causal. In particular, if individuals tend to send more applications to jobs they are more likely to get due to e.g. differences in unobservables, we would expect β_1^y to overstate the returns to search.[72] Absent a source of exogenous variation in application behavior, we have no way of removing this potential bias. To the extent that the resulting bias is relatively constant across job types, however, the estimates from Equation 10 may still give a useful ranking of the types of jobs where men face relatively higher returns to search than women.

In Figure C.9 we examine whether differences in the returns to applications appear to explain the observed gender gaps in applications. Each data point in the figure corresponds to a job type, defined as either a two-digit occupation, a two-digit industry, a decile of the firm wage level distribution or a decile of the typical wage distribution. The y-axis shows the gender gap in applications to the different job types after conditioning on labor market observables. Finally, the

[72]Since we estimate Equation 10 on the reweighted sample, we are ensuring that labor market observables are balanced across men and women. There may of course still be unobserved differences across individuals that affect the likelihood of a successful application and likely correlates with application behavior as well.

x-axis contains our (standardized) measure of the gender gap in returns to applications for the different job types.[73] If the observed gender gaps in applications were driven by women applying more to jobs where their applications have a higher relative likelihood of turning into a hire, we would expect to see an upward sloping relationship in Figure C.9. This is not what we see. If anything the relationship seems to be slightly downward sloping.[74]

The fact that gender gaps in applications are not positively correlated with gender gaps in the likelihood that an application turns into a hire is also borne out in the relative magnitudes of the application and hiring gaps documented in Section 4. If the jobs that women apply less to than men are also the ones where they face a lower chance of being hired when they apply, this will compound to make overall gender gaps in hiring outcomes even larger than gender gaps in applications.[75] As shown in Section 4, however, gender gaps in applications actually tend to be larger than gender gaps in hiring.

In sum, with the data we have available, we cannot find evidence that "self-fulfilling discrimination" underlie the observed patterns of application behavior across genders in our data. As discussed, however, this does not rule out that "self-fulfilling discrimination" is at play and could be detected with different data.[76]

C.9 Gender differences in beliefs, overconfidence and risk preferences

Another possible explanation for our results is that gender differences in preferences and beliefs lead men to systematically apply for more high-paying and harder-to-get jobs than women. If men and women have different beliefs about their general labor market prospects such that men are

[73] Specifically, we standardize the measure within each category of job types (across industries, across occupations, across deciles of firm wage levels and across deciles of typical wages). Standardizing the measure jointly across all the categories or using an non-standardized version of the measure does not change the results.

[74] The correlation between the gender gap in applications and our measure of the gender gap in returns to applications for the different job types is -0.368 overall and range from -0.762 (typical wage deciles) to 0.257 (firm wage level deciles) if computed across one of the four job categories.

[75] Put differently, overall gender gaps in hiring outcomes is the product of two gaps: 1) the gap in how likely women are to apply for a particular type of job and 2) the gap in how likely an application from a woman is to result in a hire for a particular type of job. If the types of jobs where gap 1 is big are also the ones where gap 2 is big, the aggregate gap in hiring outcomes should be bigger than the application gap alone (gap 1).

[76] As noted, our simple regression-based measure of returns to applications does not account for unobservable characteristics that correlate with both application behavior and the likeihood of getting hired. In addition, the most extreme models of "self-fulfilling discrimination", can imply that women *never* apply to any jobs where they face a lower probability being hired than men do. In this case it would never be possible to measure any meaningful differences in the likelihood of being hired conditional on applying. In our data, women are of course sending many applications to each of the observed job types we consider, however, this does not rule out that men and women are differentially applying to jobs with different unobserved characteristics within these job types.

more (over)confident than women, this could lead men to systematically target more higher-paying but harder-to-get jobs.[77] Similar predictions also arise if men are less risk-averse than women or if some form of social norms lead women to hold themselves to a higher standard when deciding where to apply for jobs. The existence of these types of gender differences have received significant attention in previous work and have also found empirical support in some settings (see in particular Cortes et al. (2020)).

To provide some evidence on this possible mechanism, we can examine the speed with which men and women find jobs in our data. To this end, we construct a version of our main analysis sample where we drop the restriction that individuals must find employment within a year (see Section 3.1).[78] For this sample we then examine how many men and women still have not found employment after each week of their unemployment spell. Specifically, Figure C.10 shows Kaplan-Meier estimates of the survivor function in nonemployment, estimated separately for men and women. To account for differences in observable characteristics of men and women, the survivor cuves are estimated after reweighting on observables (see Section 3.4).[79]

If simple gender differences in beliefs or preferences are causing women to apply to less ambitious and easier-to-get jobs than men, we should see women find jobs faster than men. Looking at Figure C.10, however, this is not what we see in the data. Throughout the unemployment spell, the survivor curve for women is slightly above the survivor curve for men, implying that if anything women in fact find jobs slightly more slowly than men do.

In sum, we cannot find evidence that simple gender differences in beliefs or risk preferences underlie the observed patterns of application behavior across genders in our data. Of course, this not rule out that gender differences in beliefs or risk preferences exist and contribute to the observed application behavior - if men and women differ in other dimensions that influence job finding rates, for example, this may obscure the effect of beliefs or risk preferences on job finding rates.

[77]It is commonplace to refer to this explanation for gender differences as reflecting male overconfidence. Of course, it is in principle also possible that men have approximately correct or even downward biased beliefs about their labor market prospects, while women are (more) underconfident and have (more) downward biased beliefs about their labor market prospects. This could also explain the gender application gaps that we see in our data through exactly the same mechanisms.

[78]We find a similar pattern if we instead focus on the main analysis sample including the restriction that individuals must find employment within a year.

[79]Figure A.1 shows corresponding estimates without reweighting. This leads to a larger difference between men and women but still with women finding jobs slower than men.

C.10 Gender-shares at applied-for firms (preferences for diversity)

In Figure 5, we focus on a set of non-wage job amenities that have been emphasized in previous work on gender gaps. Of course, there could be many other non-wage job amenities that men and women value differently. One salient example here is the gender composition of the workforce. In particular, we might think that women prefer working in more gender diverse workplaces, or alternatively that they simply prefer working with many female coworkers.

In Figure C.11 we provide some evidence in this regard. For each applied-for job, we compute the share of female employees at the firm. We then split firms into deciles of this measure and compute gender application and hiring gaps (after conditioning on observables using reweighting). As the figure shows, women are systematically more likely to apply and get hired at firms with more female employees. Moreover, the relationship between hiring/application gaps and female share is completely monotonic. If we interpret this as reflecting preferences, it would thus suggest that female UI seekers simply prefer working at firms with more female employees, not at more diverse workplace (the median firm has a female share of 33 percent).

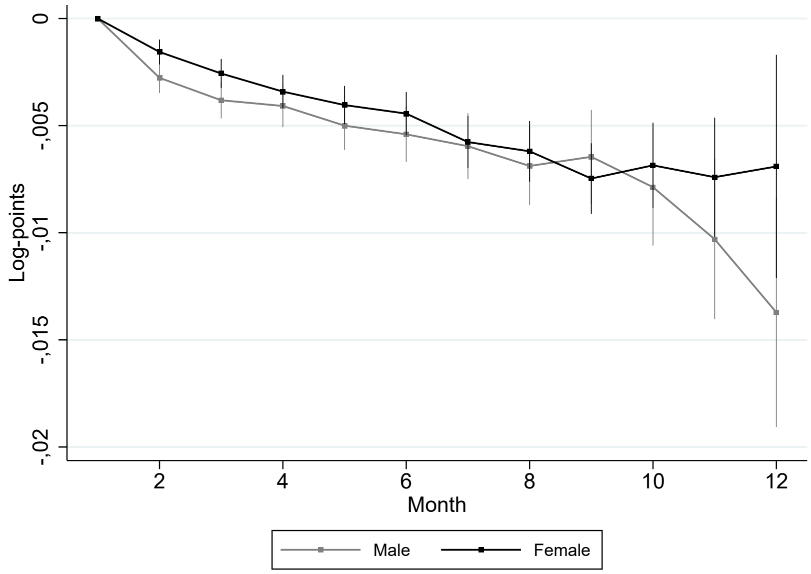
We note, however, that the figure is also consistent a situation in which women do not care about workforce composition but simply for some other reason prefer working at certain firms. Absent very particular turnover patterns, these firms should end up with more female employees in steady state, which would generate the pattern in Figure C.11.

Table C.1: Mapping job titles from Ahmed et al. (2021) to occupations

Occupation Group:	Job Title:	Observations:
Craft and Related Trades	*Vehicle mechanic*	214
Elementary Occupations	*Cleaner*	434
	Warehouse worker	141
Plant and Machinery Operators	*Truck driver*	337
Professionals	*Enrolled nurse*	206
	IT developer	153
	Pre-school teacher	270
Service and Sales	*Chef*	392
	Customer service	51
	Store clerk	127
	Telemarketing	43
	Waitstaff	497
Umatched	*Accounting clerk*	166
	B2B sales	152
	Childcare	71

Notes: The table shows our mapping of job titles in the audit study data of Ahmed et al. (2021) into 1-digit occupation groups in our data. The last column shows the number of applications sent to each job title in the audit study data.

Figure C.1: Change in job search over time, event study estimates



Notes: The figure plots estimates from an event study regression that regresses the average log typical wage of applied-for jobs on dummies for the number of months since entering unemployment. The figure plots the estimated coefficients on the month dummies. The regression includes a person-by-spell fixed effect and uses the month of entry (month 0) as the omitted baseline month. Results are shown separetely for men and women. Standard errors are clustered at the spell level.

Figure C.3: Gender differences in applications up and down the occupational job ladder

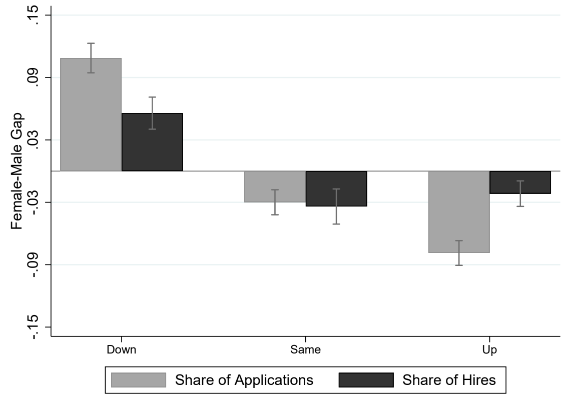
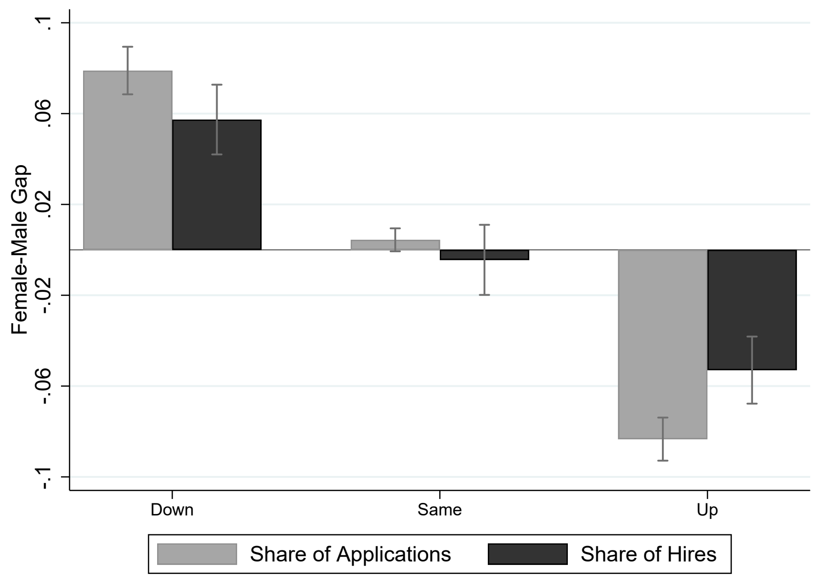


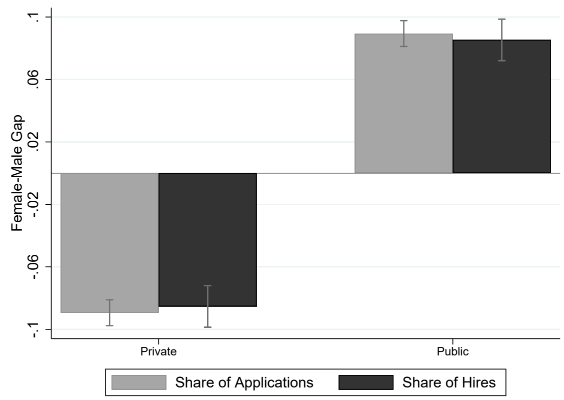
Figure C.2: Gender gaps in job applications and hiring, typical wage ranks relative to previous job



Notes: The figures plot gender gaps in the share of applications going to jobs that are in a higher, lower or the same decile as the previous job. The figure also plots corresponding gender gaps in hiring shares. All gaps are computed after propensity score reweigthing so are conditional on observables. The 95 percent confidence bars are based on standard errors clustered on the individual level.

Figure C.4: Gender gaps in job applications and hiring, public vs private sector

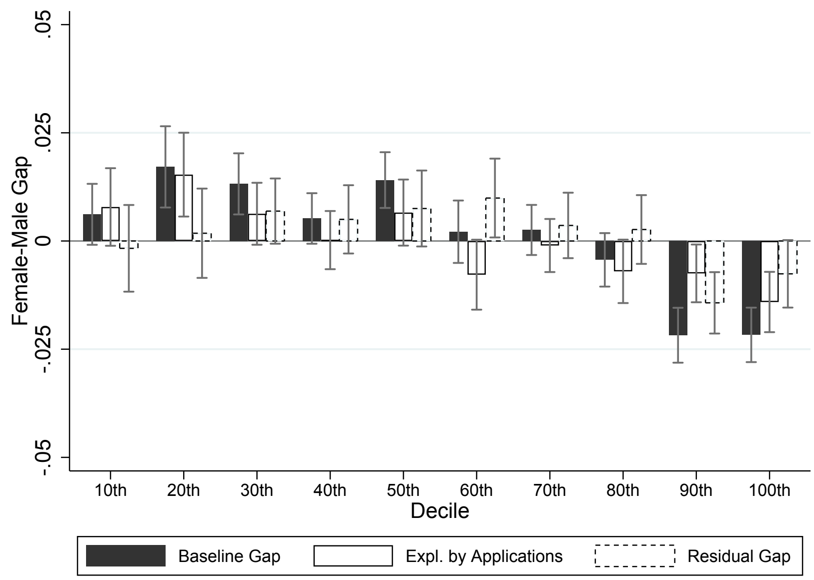
(a) Public Private



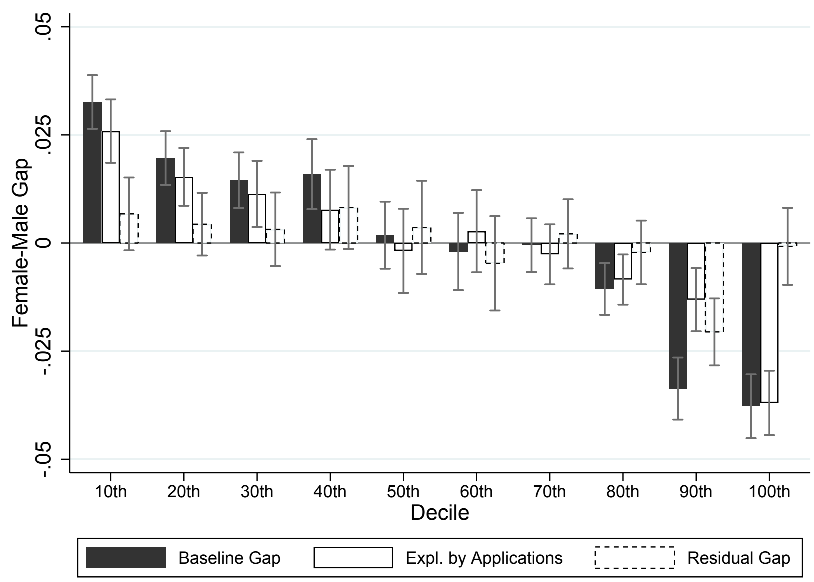
Note: Figure plots gender gaps in the share of applications going to public or private firms. We also plot corresponding hiring shares. The 95 percent confidence bars are based on standard errors clustered on the individual level.

Figure C.5: Decomposing hiring share gaps across deciles of firm types and wages

(a) Firm wage level



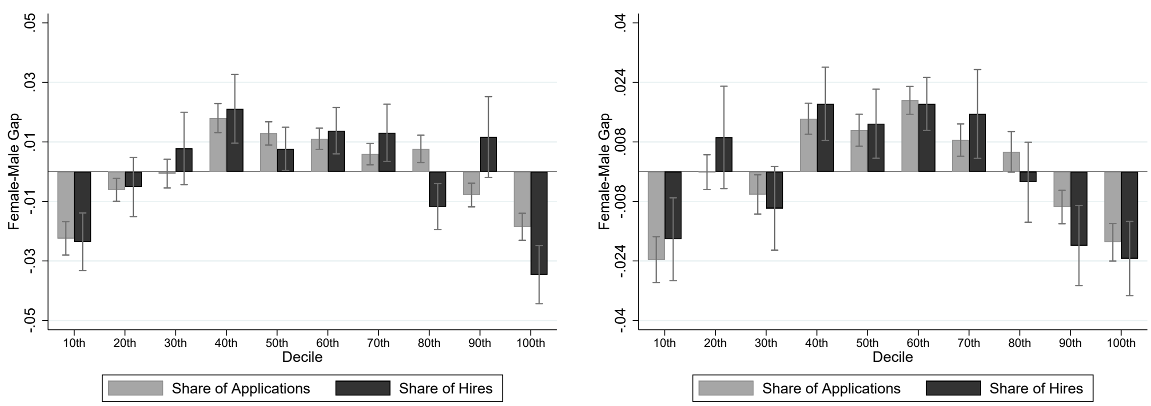
(b) Typical wages



Note: Figure decomposes baseline gaps in the share of men and women hired into different types of jobs after conditioning on labor market observables. The gaps are decomposed into a part that explained by applications and a residual gap (see Equation 5).

Figure C.6: Gender gaps in job applications and hiring, wage growth

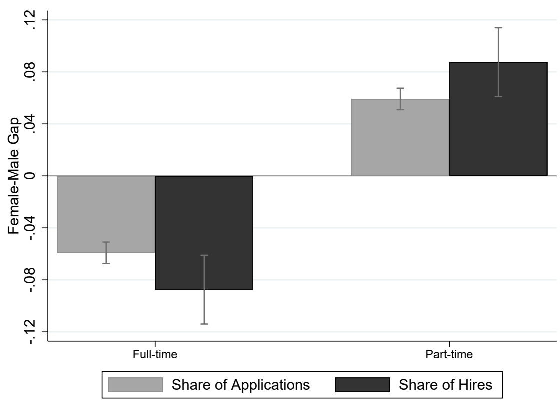
(a) Wage growth (1 year) (b) Wage growth (5 year)



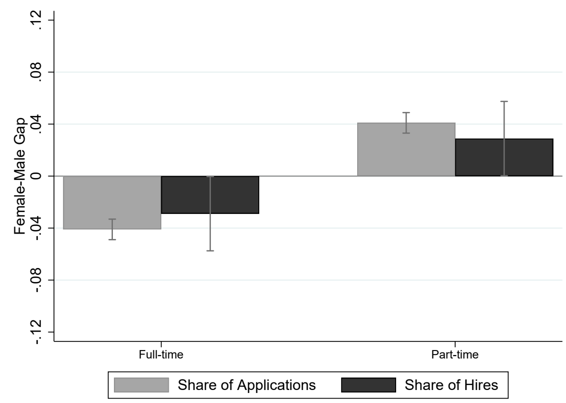
Note: Figure plots gender gaps in shares of applications going to specific wage growth deciles and gaps in which decile job-seekers are hired. Wage deciles are computed as the relative difference between the starting wage with the wage one year (left) or five years (right) after entering the respective job. The 95 percent confidence bars are based on standard errors clustered on the individual level.

Figure C.7: Gender gaps in applications and hiring outcomes, individuals with/without young children I

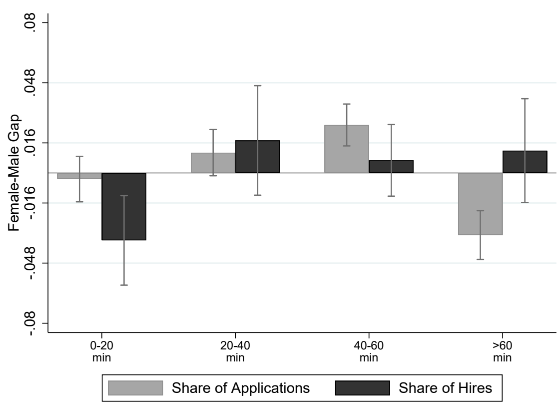
(a) Hours, individuals with young children



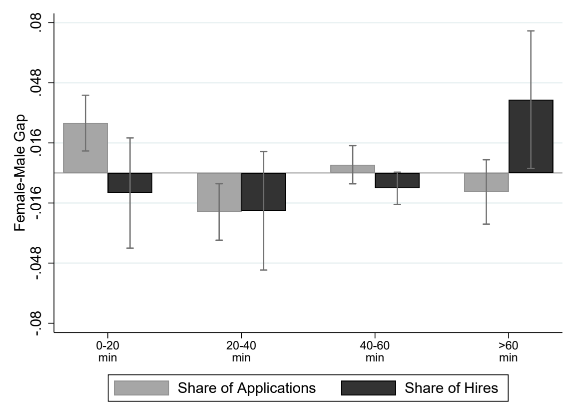
(b) Hours, individuals without young children



(c) Commuting, individuals with young children



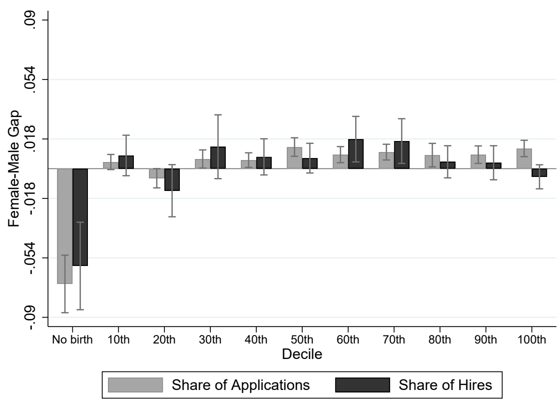
(d) Commuting, individuals without young children



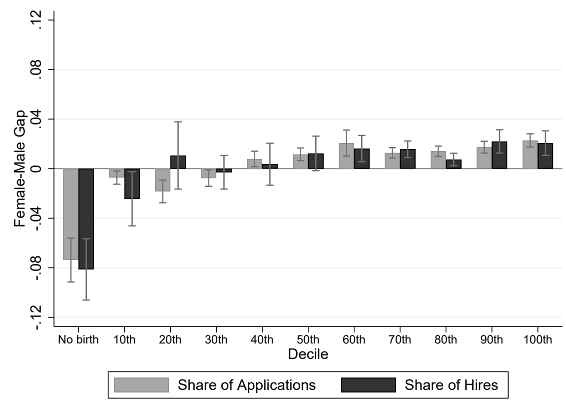
The figures plot gender gaps in shares of applications going to specific types of jobs and corresponding gender gaps in hiring outcomes separately for individuals (age 25-40) with young (0-5 years) children and without children. All gaps are based on the reweighted sample so are conditional on observables. The 95 percent confidence bars are based on standard errors clustered on the individual level.

Figure C.8: Gender gaps in applications and hiring outcomes, individuals with/without young children II

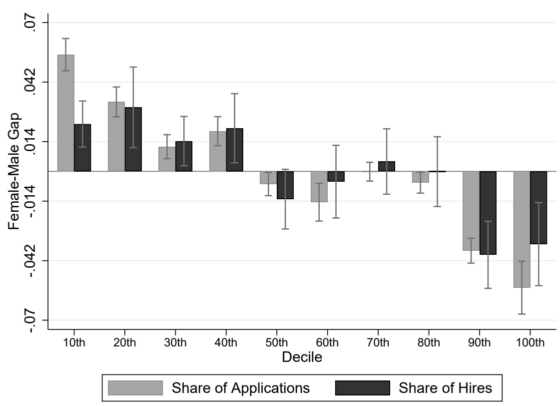
(a) Family friendly, individuals with young children



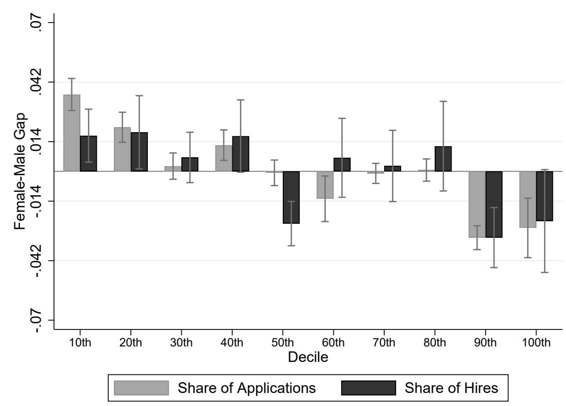
(b) Family friendly, individuals without young children



(c) Typical wages, individuals with young children

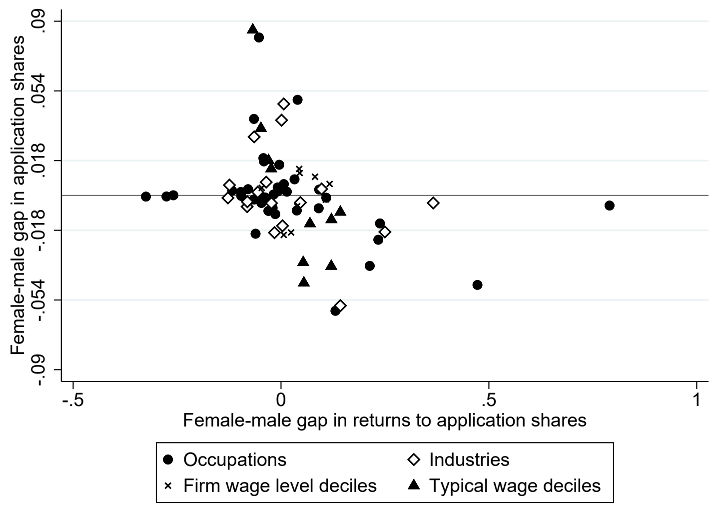


(d) Typical wages, individuals without young children



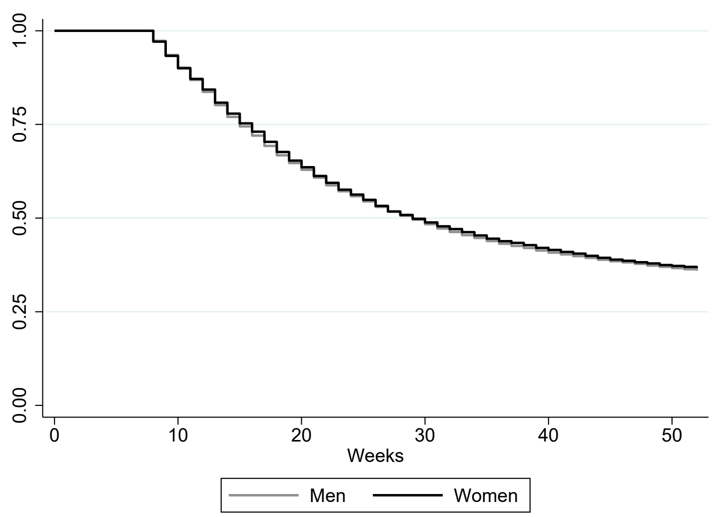
The figures plot gender gaps in shares of applications going to specific types of jobs and corresponding gender gaps in hiring outcomes separately for individuals (age 25-40) with young (0-5 years) children and without children. See Appendix A.8.1 for details on the measure of family friendliness. All gaps are based on the reweighted sample so are conditional on observables. The 95 percent confidence bars are based on standard errors clustered on the individual level.

Figure C.9: Gender application gaps and the return to applications



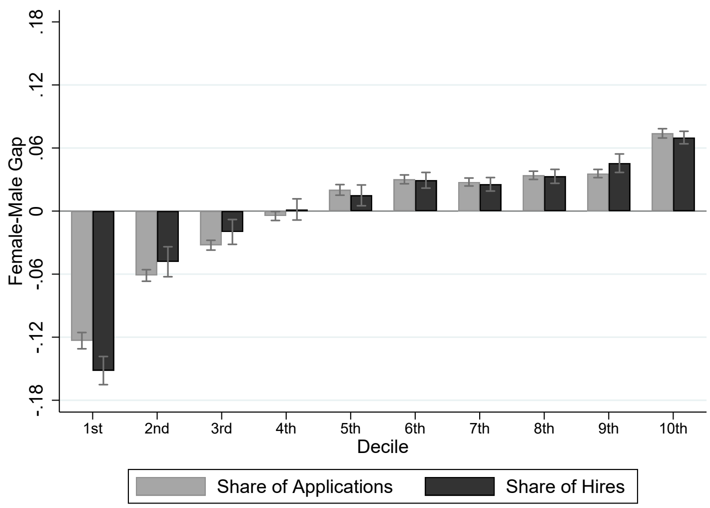
Note: The figure plots the corresponding gender application gap and the gender gap in the return to applications. Each data point is thus a job type defined as either a two-digit occupation, a two-digit industry, a decile of the firm wage level distribution or a decile of the typical wage distribution. Military occupations and job types, where fewer than 100 individuals finds employment are excluded. The y-axis shows the gender gap in applications to the different job types after conditioning on labor market observables. The x-axis contains our (standardized) measure of the gender gap in returns to applications for the different job types, see Equation 10.

Figure C.10: Kaplan-Meier survivor functions in nonemployment



Note: The figure plots Kaplan-Meier estimates of the survivor function in nonemployment, estimated separately for men and women. The two curves thus show the share of men and women that still have not found a new job after each week since starting their UI spell. The curves are estimated on a version of the main analysis sample where the requirement of finding a job within a year has not been imposed. The curves are estimated after reweigthing on observables.

Figure C.11: Gender gaps in job applications and hiring, share of females at firm



Notes: The figures groups firms into deciles in terms of the share of women in their workforce (i.e. the 10th decile contains the firms with the highest share of female employees) and then plots gender gaps in the share of applications going to each of these deciles . The figure also plots corresponding gender gaps in hiring shares. All gaps are computed after propensity score reweigthing so are conditional on observables. Occupations and industries on the x-axis are ordered from lowest paying on the left to highest paying on the right. The 95 percent confidence bars are based on standard errors clustered on the individual level.