Matching job applicants to vacancies: An empirical model

Preliminary: do not cite without author's permission

Rebecca Lessem and Robert A. Miller
Tepper School of Business, Carnegie Mellon University*

March 14, 2025

Abstract

We use data on all job applications and labor market outcomes within a large firm over a 5 year period to learn about how the search and matching process affects worker outcomes. To do this, we develop and estimate a two sided search and matching model, in which positions become vacant when the current occupant of the job leaves, the firm begins a search process by advertising the position, and workers employed both inside and outside the organization apply for the newly vacated position. Hiring is multistage, where various employees with differing objectives cull applicants through a process that leads the stakeholders to become more informed about the potential job matches. After estimating the model, we use counterfactuals to understand how multistage choice affects the selection of the hired worker from a given set of applicants, and explore how outcomes would vary with other mechanisms. We also use the model to explore how differences in hiring outcomes across racial groups and gender would be impacted under different hiring rules.

1 Introduction

There is an extensive literature on search and matching in the labor market.[1] In these papers, the process by which workers are matched with firms is often treated as a black box, primarily due to the lack of data on who applies for and ultimately receives a given job. This limitation has hindered our understanding of how firms make hiring decisions. After receiving a set of applications, how does a firm decide whom to hire? Is there evidence of competing interests within firms during this decision-making process? On the job seeker's

*rlessem@andrew.cmu.edu, ramiller@andrew.cmu.edu

[1]There are many papers in this area; see Jovanovic (1979), Pissarides (1985), Mortensen and Pissarides (1994), Burdett and Mortensen (1998), Postel-Vinay and Robin (2002), and Shimer (2005), among many others.

side, our understanding of how workers search for jobs remains limited. Traditionally, search intensity is inferred from job-to-job transitions, which is influenced by both the intensity of the search and the likelihood of receiving a job offer.[2] If we could observe job applications along with their outcomes, we could gain deeper insights into these distinct factors.

In this paper, we leverage data on job applications from a large firm to develop and estimate a search and matching model. Our primary focus is to understand the candidate selection process of the firm. In our setting, the hiring process unfolds in stages: initially, a subset of candidates is chosen for interviews; during the interview, the firm gains more insight into each interviewed candidate; subsequently, one of these interviewed candidates is offered the position. This *multistage* selection approach reduces costs compared to interviewing all applicants, but results in limited information acquisition since the firm does not gather complete information on all applicants. This work builds on a small but growing body of literature addressing hiring processes; however, to the best of our knowledge, no prior studies have utilized data on firms' actual interview selections. The closest is Gottardi et al. (2025), who develop and calibrate a model in which companies choose who to interview to understand the impact of hiring regulations. Vohra and Yoder (2024) develop a theoretical model of interview choice, albeit from a different angle, considering inter-firm competition where a firm may avoid interviewing a candidate likely to receive an offer from a more desirable position. Estimating a model of the firm's selection problem has been challenging in previous work for two reasons, both of which we overcome. First, we need data on the firm's selection process, which we have acquired for this project. Additionally, the choice of the interview set leads to a complex computational problem, as the theory suggests the firm compares sets of candidates rather than evaluating the value of each candidate individually. This results in a large choice set, posing a computational challenge that we address with our estimation strategy.

To learn more about the search and matching process, we utilize a unique dataset that provides insights into the firm's internal mechanisms. Over a five-year period in the early 21st century, we observe all job vacancies and applications within one firm. We track the hiring process through initial screening, candidate selection for interviews, and final job offers. The data reveal intriguing demographic variations on both sides of the market. African Americans and women engage in more proactive job search activities within the organization compared to Caucasian males. Black candidates are less likely to be interviewed, but when they do secure an interview, they are more likely to receive a job offer. Furthermore, post-hire outcomes differ, with both black and female employees experiencing shorter tenure and

[2]One notable exception is Faberman et al. (2022), which collects data on search intensity of workers and studies how it affects employment outcomes.

slower wage growth over their careers.

To explain these empirical observations, we develop a model in which workers choose their search intensity, positions become vacant when the current occupant leaves, and workers both inside and outside the organization apply for these newly vacated positions. The firm reviews the applications and decides whom to hire. Following the institutional rules of this firm, the hiring process occurs in two main stages: (1) an interview committee selects the candidates to interview, and (2) the firm chooses which of the interviewed candidates to hire. To allow the impact of demographic characteristics to vary at the interview and hiring stages, as observed in the data, we allow for the possibility that the interview committee's preferences do not entirely align with the firm. This approach introduces an agency problem, as the hiring committee must consider the firm's preferences when selecting interviewees, thereby adjusting the interview set based on the likelihood that their preferred candidate will be hired. The optimal interview decision involves evaluating all possible interview sets and selecting the one that maximizes expected utility. Once the interview set is chosen, the manager selects whom to hire. Wages are determined in equilibrium, reflecting the competition for the position.

We estimate the parameters of the firm's hiring decision process, which we then use to evaluate several counterfactual scenarios. First, we investigate the impact on outcomes if the firm made all hiring decisions, thereby eliminating the divergent preferences between the committee and the manager. Next, we examine the role of multistage choice by comparing outcomes in alternative settings, such as random interview selection and sequential interviews. The next set of counterfactuals explores the effects of regulatory changes on the hiring process. For example, we consider a policy that mandates the proportion of candidates interviewed from a particular gender or racial group to match their share of applications. This policy is similar to the Rooney Rule in the NFL, which requires that a minority candidate be interviewed for every head coaching vacancy. We assess how this type of regulation would influence job offer rates, application behavior, and wages over the course of a career. Additionally, we analyze the impact of the interview committee being blind to the race and gender of candidates when making interview selections.

Our research fits into multiple strands of the literature. The search and matching literature is the closest to our model and empirical framework. This is a very large literature, but some of the most relevant frameworks are in Jovanovic (1979), Pissarides (1985), Mortensen and Pissarides (1994), Burdett and Mortensen (1998), Postel-Vinay and Robin (2002), Shimer (2005), and Lentz et al. (2023). We build on this literature by modeling the worker search process as well as developing a framework for how firm's decide whom to hire from a given set of applications.

One key contribution of our work is that we identify the optimal interview set for a given application pool. This ties into the literature that aims to identify the consideration sets that agents consider when facing a large set of options. Barseghyan et al. (2021) develop a methodology to use a discrete choice model when the actual choice set is not observed, and Coughlin (2023) applies this to the choice of health insurance plans. In our work, we also estimate consideration sets, which are interview sets in our context. As compared to the prior literature, we observe the consideration sets instead of having to estimate the choice set.

After estimating our model, one of our counterfactuals examines the impact of different hiring regulations on the outcomes of minority and female candidates. This is motivated by empirical evidence in our data that shows different search behavior and job outcomes for men as compared to women, as well as for black compared to non-black candidates. It is widely known that there is a wage gap between white and black workers, as well as between men and women; in this paper we focus on understanding the role of the search and matching process in explaining this gap.[3] One potential factor that could affect these wage differentials is the likelihood of receiving a job, as well as a promotion, as this is a key determinant of wage growth over a career. Using reduced-form methods, Bertrand and Mullainathan (2004) and Shukla (2024) show that race or caste affects the probability of success after applying for a position, demonstrating how racial factors affect the demand side of the market. Russo and van Ommeren (1998), using a sample of Dutch firms, do not find any difference in the hiring rates of men and women. Babcock and Laschever (2007) shows supply side differences by documenting different negotiation strategies for men and women. Our work combines both the supply and demand sides to understand how changes in hiring regulations would affect hiring rates and life cycle outcomes.

One of the limitations of our empirical work is that we only have data from one firm, meaning that we can only learn about behavior for that firm and its applicants and employees. There is no generalizable dataset with this type of information, which means that to study the search and matching process in the way we do, we are restricted to using data from one firm. There is a substantial literature in economics that uses data from only one firm which has enabled authors to answer questions that we could not study otherwise. Examples of work that uses data from a single firm include Lazear (2000), Lazear et al. (2015), and Lazear et al. (2016).

[3]Altonji and Blank (1999), O'Neill (1970), O'Neill (1990), Neal and Johnson (1996), Carneiro et al. (2005), and Golan et al. (2024) (amongst many others) provide empirical evidence for the existence of these racial wage gaps and additionally explores some of the potential mechanisms. Research on the gender wage gap is surveyed in Blau and Kahn (2017). Other papers (amongst many others) that study the gender wage gap include Gayle and Golan (2012), Gayle et al. (2012), and Xiao (2021).

2 Data

We use a dataset with information on job applicants provided by a large anonymous firm with approximately 65,000 workers. Over a 5-year period in the first decade of the 21st century, we have information on every job application received by this firm. We know basic demographic information for each applicant (race, gender, age)[4] as well as the outcome of their application process, which we divide into stages. First, some candidates lose interest in the position, which we label as *not interested*. We mark candidates as not interested if the data record their outcome as "Applicant not interested", "Failed to respond to HR," or other similar labels. Next, some applicants may not have the required qualifications for the position. The firm labels these candidates as *not qualified*. Taking the remaining candidates who are both qualified and interested in the position, the data record if a person is interviewed for the position, if they are offered the position, and if they accept the position. The data also include basic information about the job, including the division and a short job description. The dataset provides job and candidate identifiers, which allows us to see who applies for and receives each position, and additionally track applicants' behavior across all job applications in this 5-year time period. Crucial for our analysis is that we have all the information included in the job application, which means that we have the same set of information as the hiring manager.

We supplement these data with a secondary data source, which reports annual wages for each person who works for this organization for more than a 20-year period. After merging the application with the wage data, we can observe wage outcomes after a person is hired, providing new information about the determinants of wage outcomes. We also use the wage data to learn how long a person stays at their job and at the firm.[5]

The top section of Table 1 shows some summary statistics on the sample, using an applicant as the unit of observation. During this 5-year period, there are 62,405 applicants in the data; of these, about 4.5% are African-American and a little over half are female. The average person applies for almost 4 jobs over the 5 years for which we have data. The middle section of the table uses a job as the unit of analysis. The data include information on over 5,000 vacancies, where the average job receives almost 50 applicants and interviews about 7 people. The bottom section of the table, which uses an application as the unit of analysis, shows that over 90% of applicants are external to the firm, meaning that they are not working for this firm when they submit an application. This will be an important distinction in our analysis, as we expect that internal and external applicants may be treated

[4] Age is reported in bins.

[5] In the wage data, the division is recorded for each year. We assume that a person moves jobs when they switch divisions or leave the firm.

differently in the application process.

2.1 Search behavior

We analyze the search behavior of workers to understand how these decisions are made in order to motivate the model. As a first component of job search behavior, we analyze the determinants of the number of positions a person applies to in a given year. Table 2 shows the results of a regression where the dependent variable is the number of applications that a person submits in a year. The results show that women and African-American candidates apply for more positions and people with more experience in this firm apply for more jobs. The second column only uses applications where the candidates are *qualified* and *interested* in the position. These results are qualitatively similar to column (1).

Next, we look at the type of positions that people apply for, which we measure as the pay level of the job. Table 3 shows the result of a regression where the dependent variable is the salary level of the job that a person applies for.[6] Columns (1) and (2) show all applications, and columns (3) and (4) show only applications where the candidate is qualified and interested. Looking at the results in columns (1) and (3), we see that women and African-Americans apply to lower salary positions. Columns (2) and (4) additionally include controls for a person's prior year salary. These sample sizes are much smaller since we only know prior salary for candidates who are working at the firm when they apply for a new position. After controlling for prior-year salary, the result that African Americans and women apply for lower-salary jobs holds (when looking at applications where the candidate is qualified and interested).

Our data record when a candidate begins an application, and in many cases (14%) the applicant does not complete the application, which is labeled as *not interested* in the data. We interpret this as the candidate learning more about the position and deciding it is not a good fit for them. We run a probit regression to understand the determinants of a candidate being uninterested in a position. The results, shown in column (1) of Table 4, indicate that African Americans and women are more likely to apply for positions they are not interested in. We will use this in the model to learn about people's preferences over different types of jobs.

[6]The firm provides the minimum and maximum salaries, and we take the average to use our analysis. If we were to use the minimum or maximum values, the results of this analysis are very similar.

2.2 Application review process

The application review process occurs as follows. First, applicants are screened for minimal qualifications, which we denote as *qualified*. Next, a group of candidates is selected for an interview, and then one of them receives a job offer. The person who receives the job offer can accept or reject the offer.

We run a series of probit regressions to understand the application process, looking at whether the candidate is qualified, interviewed, and then hired. Each of these regressions controls for race, gender, experience in the firm and division, and the salary of the job posting. We also include year, division, and occupation fixed effects. Tables 4, 5, and 6 show the probit regression results for being not qualified, interviewed, and hired, respectively.

The second column of Table 4 shows the results of a probit regression where the outcome variable equals 1 if a person is not qualified for a position, which happens 15% of the time across all applications. The regression results show that African Americans are more likely to apply for positions they are not qualified for.

In Table 5, the outcome variable equals 1 if a qualified and interested candidate is interviewed for a position. Column (1) uses all candidates, and column (2) uses applicants for jobs that only hired one person. African Americans are less likely to be interviewed. In rows columns (3) and (4), the sample is split based on whether or not the candidate has prior experience working at this firm, to examine whether the coefficients on race and gender vary for candidates that the hiring committee may personally know. This does not affect the qualitative results.

Table 6 uses an outcome variable that equals 1 if an interviewed candidate is offered the position. Column (1) uses all applications, column (2) only uses positions that hire one person, and columns (3) and (4) split the sample based on whether or not the candidate has prior experience at the firm. Although African-American candidates are less likely to be interviewed, these results show that, once they get an interview, they are more likely to be hired. This shows a potentially different decision process at these two stages that we will account for in the model.

We have done a similar exercise on the likelihood that a person accepts a job offer. However, the data do not show any interesting or significant trends in this analysis, which is likely because more than 95% of the job offers are accepted. For these reasons, we do not report these regression results.

2.3 Outcomes on the job

This subsection analyzes two outcomes on the job: how long a person stays at a job and their wages while employed at this firm. Job duration could be an important component of the firm's selection process, since a manager's expectations as to how long a person will remain at a job affects hiring decisions. Table 7 explores the relationship between the duration a person remains at a job and their race, gender, and other personal characteristics as well as the characteristics of the job.[7] Column (1) shows OLS results, and column (2) shows the results of a Tobit, which controls for the truncation that occurs since many people are still at their job at the end of the sample period.[8] These results show that both African Americans and women stay at jobs for shorter periods of time. This result was also found by Gayle et al. (2012), which looks at a very different sample consisting of executives from large firms.

We estimate a wage regression to examine the relationship between demographic characteristics, experience, and wages. Although the data include all people in the firm, we only include people who apply for a job in the application sample period, since that is the only group for which we have race information. For this sample of individuals, we have wage observations over many years. Table 8 shows the result of regressions where the dependent variable is log earnings each year. Column (1) uses all observations and shows that African Americans earn lower wages, and men earn higher wages. Among other things, we also control for whether a person switched divisions in a given year, which is a strong signal that they switched jobs; this coefficient is positive, indicating that job switches lead to higher earnings. Next, we aim to learn what happens to wages when a person switches jobs. To do this, we restrict our sample to the time period in which we observe applications; this allows us to see if and when a person switches to a new job. These results are in column (2) and show that switching to a new job leads to higher wages. In columns (3) and (4), we restrict to our application sample period and control for the number of applications submitted by a candidate in the prior year. This allows us to examine if on the job search behavior affects wages. We see a small negative (and weakly statistically significant) relationship between the number of applications submitted and wages, and no statistically significant relationship when we only consider the qualified and interested applications. The model of Postel-Vinay and Robin (2002) predicts that job search leads to increased wages due to bargaining between the new job and the original employer. In this setting, our data reports job search directly,

[7] To calculate durations, we look at the number of consecutive years a person works in a given division.

[8] Here is an explanation of why we additionally control for the year the job started in the Tobit model. Start with the regression model $y = X\beta + \gamma t_0 + \varepsilon$, where y is the duration of the job, X is characteristics, and t_0 is the year the job started. We do not see y for all observations; instead we see y^*, where $y^* = \min\{y, T_f - t_0\}$, where T_f is the final year for which we have data. We can rewrite this as $y^* + t_0 = \min\{y + t_0, T_f\}$. Then the regression is specified as $y^* + t_0 = X\beta + (\gamma + 1)\, t_0 + \varepsilon$.

and somewhat surprisingly, we do not see evidence supporting this positive relationship.

Table 9 further restricts the sample to people who change jobs. These results do not show a statistically significant relationship between race and wages (although this is possibly due to the small number of minority workers in the smaller sample used to estimate this regression). People who switch divisions earn more, and this effect is significant at a 10% significance level. There is a negative association between the number of applicants for a position and a person's wage in that role, suggesting a role for compensating variations for more popular jobs. Column (2) additionally controls for the salary of the job posting, and most of the prior results hold. Interestingly, women earn lower wages, even when conditioning for the salary of the position. Recall that earlier tables showed that women apply for lower paying jobs on average. This gives two reasons for the gender wage gap: (1) job choice and (2) wage outcomes conditional on the job level. The result that women earn less, even after conditioning on the pay range, can be explained by two mechanisms. The first is that women negotiate for nonpecuniary amenities, such as a flexible work schedule. The difference could also be due to variations in the negotiation effort, as in Babcock and Laschever (2007). Column (3) continues to use all job switchers and analyzes the wages in all years after their job change. We interact the salary of the job posting with the years since the person started a new job.

One limitation of our analysis is that our data only contain information from one firm, which may not be representative of the region where it is located. To assess this problem, we computed statistics from the ACS to compare to our data. This comparison in shown in Appendix A.

3 Model

This section develops a continuous time model of a multistage hiring process within a stationary economy, where production is additively independent across jobs that confer financial compensation and nonpecuniary benefits. Our analysis is organized around the three types of decisions, a manager's choice over interviewed applicants to fill a job vacancy, a hiring committee's selection about whom to interview, and a worker's decision about which job applications to pursue and the terms of employment he would accept. A pool of minimally qualified workers denoted by $\mathbb{V}$ are presented with the opportunity to apply for a newly vacated position. A subset of them denoted by $\mathbb{A} \subseteq \mathbb{V}$ apply for the job. Then a hiring committee vets each applicant $a \in \mathbb{A}$ and selects a consideration set $\mathbb{C} \subseteq \mathbb{A}$ for interview. Each candidate $c \in \mathbb{C}$ selected for interview forms a reservation wage for relocating, which plays a role in final stage of the hiring process, where the manager makes as many job offers

as is necessary to fill the position.

Workers New jobs start at discrete intervals in time when the worker quits a previous position. We label job spells by $i \in \{1, 2, \ldots\}$, denoting by τ_i the time that the worker completes the i^{th} spell. Jobs offer workers wages and nonpecuniary benefits. When a worker accepts a new position he incurs an initial adjustment or job relocation cost, which we denote by ϵ_i. We assume ϵ_i is independently and identically distributed over spells. Over the duration of the i^{th} job spell the worker is compensated by a wage flow rate of w_i, incurring nonpecuniary benefits and costs at the rate of u_i.[9] The expected current value of pecuniary and nopecuniary benefits for the duration of the i^{th} job spell is thus $\delta^{-1} E_{\tau_{i-1}} \left[e^{\delta(\tau_{i-1} - \tau_i)} \right] (w_i + u_i)$ where the expectations operator $E_t [\cdot]$ conditions on information a can infer at time t.

The nonpecuniary components also include the costs of pursuing other employment opportunities on the job. Thus wages and nonpecuniary benefits are endogenously determined by how aggressively workers search for new employment. Employment opportunities occur at random times denoted by ρ_{ij} where $j \in \{1, 2, \ldots\}$ indicates the j^{th} new employment opportunity within the i^{th} job spell. At ρ_{ij} the worker chooses whether to pursue the opportunity by setting $d_{ij} = 1$, or decline it by setting $d_{ij} = 0$. If $d_{ij} = 1$ he incurs submission costs of ξ_{ij}. Setting $d_{ij} = 1$ and incurring the submission cost is a necessary condition for filling the vacancy but not guarantee success. Applicants compete for positions by optimally indicating their enthusiasm for the job if they are selected into the consideration set for interview. Let J_i denote the particular application that leads to promotion from the i^{th} position at time τ_i. Without loss of generality set the current time to $t = 0$, and label the current job as the first. The worker's current expected lifetime utility from then onwards, optimally pursuing new employment opportunities, is then:

$$V = \max_{\{d_{ij}\}_{(i,j)}} E_0 \left[\sum_{i=1}^{\infty} \left\{ \begin{array}{l} \delta^{-1} e^{\delta(\tau_{i-1} - \tau_i)} (w_i + u_i) \\ + e^{-\delta \tau_i} \epsilon_{i+1} + \sum_{j=1}^{J_i} d_{ij} e^{-\delta \rho_{ij}} \xi_{ij} \end{array} \right\} \right] \tag{1}$$

where δ is the discount factor, $\tau_0 = 0$ by our timing convenstion, and the definition of J_i implies $\sum_{j=1}^{J_i} d_{ij} \exp(-\delta \rho_{ij}) \xi_{ij} = \exp(-\delta \rho_{iJ_i}) \xi_{iJ_i}$.

Suppose the worker is currently facing a new employment opportunity, and let $d \in \{0, 1\}$ denote the worker's decision to submit an application or not. Denote by $w + u$ the flow benefits within his current job, and the submission cost for the new employment opportunity by ξ. If he pursues the new employment opportunity by setting $d = 1$, his relocation cost ϵ (or alternatively intepreted, that part of the orientation cost incurred by the applicant if

[9] To simplify the notation, the model assumes wages are constant within a spell. Relaxing this assumption is straightforward, and our empirical specification allows for wage adjustments within spells.

he takes the new position),and the nonpecuniary amenity flow of the new job, denoted by $\widehat{u}$ is revealed. Let $\widehat{w}$ denote the wage paid to the successful applicant, and let $\widehat{\epsilon}$ denote his relocation cost. Thus $\widehat{\epsilon}$ is only defined for switchers, in which case $\widehat{\epsilon} = \epsilon$. Denote by $\widehat{V}$ the value of the social surplus function for the worker with a job yielding benefit flow $\widehat{w} + \widehat{u}$.

Several factors decide whether to submit an application or not for the job. The potential benefits include $\widehat{w} - w$, the difference in wages between current employment and the advertised job; $\widehat{u} - u$, the difference in their nonpecuniary benefits; $e^{-\delta\rho}$, the discount factor waiting until the next employment opportunity appears that might partially equalize these differentials; and $\widehat{V} - V$, the difference in the expected lifetime utility at that future point in time. These potential benefits are weighted by ϕ, that the worker is offered and accepts the new job. Balanced against these potential gains are $\widehat{\epsilon}$, the relocation cost of moving also weighted by the probability of being hired, and ξ, the submission cost. Lemma 1, proved in the Appendix for a more general case where wages vary with job opportunities and experience, show how the components are linked.

Lemma 1 *It is optimal for the worker to submit an application if and only if:*

$$\delta\left(\xi + \phi\widehat{\epsilon}\right) \leq \phi E\left[\left(1 - e^{-\delta\rho}\right)\left(\widehat{w} + \widehat{u} - w - u\right) + e^{-\delta\rho}\delta\left(\widehat{V} - V\right)\right] \tag{2}$$

As in much of the search literature, the reservation wage plays a role in determining who is hired and what they are paid. Denote by $\overline{w}$ the worker's reservation wage, and denote by $\overline{V}$ denote the worker's social surplus function, or expected lifetime utility upon taking the new job he is offered following a successful application. Lemma 2 shows the equilibrium relationship between the two.

Lemma 2 *The reservation wage and the continuation value for a worker who starts a job at the reservation wage are uniquely defined by the equation:*

$$\left(1 - E\left[e^{-\delta\rho}\right]\right)\left(\overline{w} + \widehat{u} - w - u\right) = \delta\left\{E\left[e^{-\delta\rho}\right]\left(V - \overline{V}\right) + \epsilon\right\} \tag{3}$$

The initial benefit flow from turnover at the reservation wage is $\overline{w} + \widehat{u} - w - u$. This benefit accrues only until the next employment opportunity appears, explaining the expression $\left(1 - E\left[e^{-\delta\rho}\right]\right)$. At the next employment opportunity, the discounted value of the difference in continuation values, reprepresenting expected liftime utility at the point in time is $E\left[e^{-\delta\rho}\right]\left(V - \overline{V}\right)$. Scaling this expression by δ calibrates it in flow terms, and likewise for ϵ, the relocation cost of switching jobs.

Manager Suppose $c_k \in \mathbb{C}$, a candidate selected by the committee from $\mathbb{A}$, the pool of applicants, is hired into an advertised position at time $t = 0$. Let π_k denote his instantaneous product flow while in this job spell, and suppose he is paid at a wage rate w_k determined upon hiring. The annuity value to the manager at the time he is hired is:

$$M(\pi_k, w_k) = \pi_k - w_k + E\left[e^{-\delta\tau_k}\right][M_0 - (\pi_k - w_k)] \tag{4}$$

where M_0 is the annuity on the social surplus value to the firm from filling the vacancy, that is the exante value of the position to the firm after it becomes vacant but before the applicant pool forms, and τ_k is the random time c_k quits this job spell. The manager's payoff from filling the vacancy, given by (4), implies she is indifferent between $c_k \in \mathbb{C}$ and $c_{k'} \in \mathbb{C}$ if and only if the compensation offered to each of them satisfies the equality $M(\pi_k, w_k) = M(\pi_{k'}, w_{k'})$. This equality specializes to $w_k - w_{k'} = \pi_k - \pi_{k'}$, the compensating differential that would arise in a static model, or more generally when their quitting times share the same probability distribution. Alternatively set $w_k - w_{k'} = \pi_k - \pi_{k'}$ but suppose $E\left\{\exp\left[\delta\left(\tau_k - \tau_{k'}\right)^-\right]\right\} > 0$. From (4), if $M(\pi_k, w_k) > M_0$ then $M(\pi_k, w_k) > M(\pi_{k'}, w_{k'})$, but if $M(\pi_k, w_k) < M_0$ then $M(\pi_k, w_k) < M(\pi_{k'}, w_{k'})$, because when the consideration set is less attractive than expected, longer spell durations count as a negative factor, but if they are more attractive than expected, longer durations counts in a positive way.

Equilibrium hiring and wages in our model is supported by several different bargaining mechanisms that are strategically equivalent.[10] To describe one of the simplest, suppose that in the final stage of the hiring the firm commits to each candidate $c_k \in \mathbb{C}$ not to offer any wage lower than the amount r_k that c_k designates at the beginning of the interview as the minimal compensation he would accept. The manager relabels the candidates so that $M(\pi_k, w_k) > M(\pi_{k+1}, w_{k+1})$ for all $k \in \{1, \ldots, |\mathbb{C}| - 1\}$, and makes an ultamatum offer to c_1 at a wage $\widehat{w}_1$ satisfying the equality $M(\pi_1, \widehat{w}_1) = M(\pi_2, r_2)$. If c_1 rejects the offer, the manager offers c_2 a wage of $\widehat{w}_2$ where $M(\pi_2, \widehat{w}_2) = M(\pi_3, r_3)$. This continues until a candidate accepts her offer, or everybody rejects their respective offers, in which case the firms pays a penalty to restart the hiring process at some future date by drawing a new set of applicants. The information the applicant reports to the manager helps determine whether he is offered the job, but not his wage conditional on being offered the job. Revealing his reservation wage is therefore a dominant strategy. Because $M(\pi_k, w)$ is decreasing in w, it

[10]For example, this mechanism yields the same outcome that would occur if applicants were tested for their enthusiasm for the job in the analogue to a decending price private value procurement auction with differential scoring to account for heterogenous quality: the personnel manager offers sucessively less attractive wages, keeping $M(\pi_c, w_c)$ equalized across the shrinking subset of applicants within the consideration set, until only one remains. In this respect there is a close parallel to the auction literature. See for example Hale and Tamer.

follows that $\widehat{w}_k > \overline{w}_k$ and hence c_k would accept the manager's offer.

Theorem 3 *There is a unique perfect equilibrium for this game. In equilibrium* $r_k = \overline{w}_k$, *defined in* (3) *for each* $k \in \{1, \ldots, |\mathbb{C}| - 1\}$, *and Let* $\overline{\tau}_k$ *denote the optimal time for* c_k *to turnover when he is paid his reservation wage after accepting the new position, and* $\widehat{\tau}_k$ *denote the optimal turnover time from the new position given the equilibium ultamatum wage* $\widehat{w}_k$ *he accepts. The equilibrium wage that would be offered to* c_k *is:*

$$\widehat{w}_k = \overline{w}_{k+1} + \pi_k - \pi_{k+1} + \frac{E\left[e^{-\delta\widehat{\tau}_k}\right] - E\left[e^{-\delta\overline{\tau}_{k+1}}\right]}{1 - E\left[e^{-\delta\widehat{\tau}_k}\right]}\left(\overline{w}_{k+1} - \pi_{k+1} + \delta M_0\right) \tag{5}$$

On the equilibrium path c_1 *accepts the manager's job offer of* $\widehat{w}_1$ *and:*

$$M\left(\pi_1, \widehat{w}_1\right) = E\left[e^{-\delta\overline{\tau}_2}\right] M_0 + \left\{1 - E\left[e^{-\delta\overline{\tau}_2}\right]\right\}\left(\pi_2 - \overline{w}_2\right)$$

The rent to the successful candidate, $\widehat{w}_1 - \overline{w}_1$ reflects differences in flow productivity and expected duration on the job. For example the equilibrium wage might be lower than the reservation wage of all the other applicants even if the flow productivity of the hired worker dominates everyone. This outcome can occur if $\pi_2 - \overline{w}_2 > \delta M_0$ but $E\left[e^{-\delta\widehat{\tau}_1}\right] > E\left[e^{-\delta\overline{\tau}_2}\right]$. loosely put the quality of the applicant pool is higher than expected but the successful applicant is expected to stay with the job longer than his closest rival, or the reverse $\delta M_0 > \pi_2 - \overline{w}_2$ but $E\left[e^{-\delta\widehat{\tau}_1}\right] < E\left[e^{-\delta\overline{\tau}_1}\right]$.

The Committee It costs η to interview each candidate if more than one is interviewed. The committee may differentially value other personal characteristics of the applicants, and denote that difference in preferences with regards to $a \in \mathbb{A}$ by $L(a)$. Conditional on $\mathbb{A}$ and any $\mathbb{C} \equiv \{a_k\}_{k=1}^{|\mathbb{C}|} \subseteq \mathbb{A}$, let $\pi_{\mathbb{C}} = \left(\pi_1, \ldots, \pi_{|\mathbb{C}|}\right)$, let $\overline{w}_{\mathbb{C}} = \left(\overline{w}_1, \ldots, \overline{w}_{|\mathbb{C}|}\right)$ and denote by $F_{\mathbb{C}}\left(\pi_{\mathbb{C}}, \overline{w}_{\mathbb{C}} \,|\mathbb{A}\right)$ the joint probability distribution of productivities and reservation wages for all the applicants in $\mathbb{C}$. Also let c_1 denote the successful candidate, and c_2 the runner up at the final stage. The committee's selection of $\mathbb{C}$ shapes the manager's choice set over the winning candidate ($a_i = c_1$ for some $a_i \in \mathbb{C}$) and the runner-up ($a_j = c_2$ for some $a_j \in \mathbb{C}$); the committee maximizes expected net productivity $M\left(\pi_1, \widehat{w}_1\right)$ augmented by the committee's own preferences $L(c_1)$ less total interview costs $\eta\,|\mathbb{C}|$, recognizing from (5) that greater competition from the runner up c_2 lowers the winner's wages $\widehat{w}_1$ and consequently increases $M\left(\pi_1, \widehat{w}_1\right)$. Since the committee has less information than the manager, it integrates over the unobserved characteristics of all possible combinations of the two top candidates $\mathbf{1}\{(a_i, a_j) = (c_1, c_2)\}$ in a given $\mathbb{C}$ using $F_c\left(\pi_c, \overline{w}_c\right)$. Formally, the committee selects $\mathbb{C} \subseteq \mathbb{A}$

to maximize:

$$\mathbf{L}(\mathbb{C}) \equiv \sum_{(a_i,a_j)\subseteq\mathbb{C}} \int \mathbf{1}\{(a_i,a_j)=(c_1,c_2)\}\left[L(a_i)+M(\pi_i,\widehat{w}_i)\right] dF_{\mathbb{C}}(\pi_{\mathbb{C}},\overline{w}_{\mathbb{C}}\,|\mathbb{A}) - \eta\,|\mathbb{C}| \tag{6}$$

In principle, the committee's optimization problem given by (6) can be solved by evaluating and comparing the $2^{|\mathbb{A}|}-1$ potential consideration sets induced by $\mathbb{A}$, but this exercise is infeasible for large $|\mathbb{A}|$. However, the number of relevant comparison sets declines to at most $|\mathbb{A}|-1$ if three additional conditions are satisfied. First assume conditional on the committee's information about $\mathbb{A}$, that $(\pi_k,\overline{w}_k)$ is independently distributed for all $a_k\in\mathbb{A}$ with probability distribution function $F_k(\pi_k,\overline{w}_k\,|\mathbb{A})$. Second, the committee can order applicants, by first order stochastic dominance (FOSD), the probability distribution that $F_k(\pi_k,\overline{w}_k\,|\mathbb{A})$ induces on to $M(\pi_k,\overline{w}_k)$. Third, the committee preferences coincides with the manager's. Theorem 2 below implies that by FOSD, every applicant in the consideration set ranks above every excluded one.

A discrete analogue to a first order condition characterizes how many applicants are included in the optimal consideration set $\mathbb{C}^o$, denoted by $|\mathbb{C}^o|$. The global optimality property of this marginal condition arises because the bargaining mechanism also solves a concave social surplus optimization problem in which $\mathbb{C}^o$ maximizes:

$$\overline{\mathbf{L}}(\mathbb{C}) \equiv \int_{\mathbb{C}} \max_{a_k\in\mathbb{C}}\{M(\pi_k,\overline{w}_k)\}\, dF_{\mathbb{C}}(\pi_{\mathbb{C}},\overline{w}_{\mathbb{C}}\,|\mathbb{A}) - \eta\,|\mathbb{C}| \tag{7}$$

over $\mathbb{C}\subseteq\mathbb{A}$ if the committee's preferences are aligned with the manager's. The key difference between $L(\mathbb{C})$ specialized to $L(a)=0$ and $\overline{L}(\mathbb{C})$ is that the former is the manager's criterion function in equilibrium, whereas the latter is a social surplus function assigning all the surplus to the firm.

Theorem 4 *Assume* $F_{\mathbb{C}}(\pi_{\mathbb{C}},\overline{w}_{\mathbb{C}}\,|\mathbb{A}) = \prod_{k=1}^{|\mathbb{C}|} F_k(\pi_k,\overline{w}_k\,|\mathbb{A})$ *and suppose there exists an ordering within* $\mathbb{B}$ *denoted by* $\{c_k\}_{k=1}^{|\mathbb{C}|}$ *such that for all* $k<|\mathbb{C}|$*:*

$$\int\int_{M(\pi_k,\overline{w}_k)\leq M} dF_k(\pi_k,\overline{w}_k\,|\mathbb{A}) \leq \int\int_{M(\pi_{k+1},\overline{w}_{k+1})\leq M} dF_{k+1}(\pi_{k+1},\overline{w}_{k+1}\,|\mathbb{A}) \tag{8}$$

Also assume $L(c_k)=0$ *and let* $\mathbb{C}_N=\{c_k\}_{k=1}^{N}$ *for* $N\leq|\mathbb{C}|$*. If* $\overline{\mathbf{L}}(\mathbb{C})\geq\overline{\mathbf{L}}\left(\mathbb{C}_{|\mathbb{A}|-1}\right)+\eta$ *then* $\mathbb{C}^o=\mathbb{A}$*. Otherwise* $\mathbb{C}^o=\mathbb{C}_N$ *where* N *is the unique solution to:*

$$\overline{\mathbf{L}}(\mathbb{C}_N) \geq \max\left\{\overline{\mathbf{L}}(\mathbb{C}_{N-1})+\eta, \overline{\mathbf{L}}(\mathbb{C}_{N+1})-\eta\right\}. \tag{9}$$

The theorem delivers a parsimonious algorithm for solving the committee's problem: follow the ranking by adding candidates to the consideration set up to the point where the marginal increment in $\overline{\mathbf{L}}(\mathbb{C}_N)$ is less than the interview cost. Because FOSD is an ordinal measure, the ranking of candidates does not uniquely define preferences. Consequently one result in the theorem can be extended by slightly relaxing the assumption of common preferences. In pariticular, if the rank ordering of the committee coincides with the manager's, then for a prespecified size in the consideration set, the committee selects the same set as the manager would. Let $\mathbb{C}_N$ denote the solution for choosing $\mathbb{C} \subseteq \mathbb{B}$ to maximize (6) subject to the constraint $|\mathbb{C}^o| \leq N$. when $\eta = 0$. However the committee might not select the same number of candidates to be interviewed, because the interview cost is cardinal, affecting how many candidates are selected. For example the committee might value the marginal candidate and those ranked below him very differently to the manager.

Corollary 5 *Suppose there exists an ordering* $\{c_k\}_{k=1}^B$ *defined by* (8)*, but that* $L(c_k) \neq 0$ *for some* $k \in \{1, \ldots, B\}$*. Then* $L(c_k) + M(\pi_k, \overline{w}_k)$ *also satisfies FOSD with an ordering* $\{c'_k\}_{k=1}^B$*. If* $c_k = c'_k$ *for all* $k \in \{1, \ldots, B\}$ *then* $\mathbb{C}'_N = \mathbb{C}_N$*.*

To illustrate how the algorithm fails if the conditions are violated, we consider a simple static environment ($\delta = 0$), where current wages are zero ($w = 0$), as are the nonpecuniary benefits from current and future work ($u = \widehat{u} = 0$), and there are no relocation costs ($\epsilon_c = 0$). In the first two examples the committee's preferences coincide with the manager's ($L(c) = 0$).

Example 6 (Relaxing independence) *The new hire will undertake one of three tasks* $j \in \{1, 2, 3\}$*, revealed to the manager after the committee selects* $\mathbb{C}$*. There are four types of applicant labelled by* $k \in \{1, \ldots, 4\}$*, and at least two of each in the applicant pool* $\mathbb{A}$*. The productivity of type* k *undertaking task* j *is denoted by* π_{kj}*. The committee views each task as equally likely. Figure 1A displays the productivity of type* k*, a triplet denoted by* $\pi_k = (\pi_{k1}, \pi_{k2}, \pi_{k3})$*. The optimal composition of consideration sets for* $N \in \{2, \ldots, 6\}$ *is*

displayed in Figure 1B, along with $L(\mathbb{C}_N)$*:*

Figure 1A

productivity k \ *task* j	1	2	3
π_1	9	9	0
π_2	0	0	9
π_3	0	12	0
π_4	12	0	0

Figure 1B

N	$\mathbb{C}^N$	$L\left(\mathbb{C}^N\right)$
2	$\{2,0,0,0\}$	6
3	$\{2,0,1,0\}$	6
4	$\{0,0,2,2\}$	8
5	$\{1,2,1,1\}$	9
6	$\{0,2,2,2\}$	11

The first two types are complementary, because $cov(\pi_1, \pi_2) = -18 < 0$, whereas the first type is is a partial subsitute for the third and fourth, because $cov(\pi_1, \pi_3) = cov(\pi_1, \pi_4) = 12 > 0$. Noting $\mathbb{C}^2$ and $\mathbb{C}^6$ do not intersect, the general purpose candidate is selected into a small consideration set when interview costs are high, but not into a larger consideration set. When interview costs are lower, the optimal consideration set comprises a larger portfolio of specialists who complement each other by spanning different tasks (the other three types when $N = 6$). Notice too that $\mathbf{L}(\mathbb{C}_2) = \mathbf{L}(\mathbb{C}_3)$ while $\mathbf{L}(\mathbb{C}_6) + \mathbf{L}(\mathbb{C}_4) > 2\mathbf{L}(\mathbb{C}_5)$: the gross value (before interview costs) is not, loosely speaking, concave in the size of the optimal consideration set[11], another irregularity that confounds optimization.

Example 7 (FOSD fails) *The productivity of the first of two types of applicant is an independent random variable uniformly distributed on* $[\underline{\pi}, \overline{\pi}]$ *where* $\underline{\pi} + \overline{\pi} > 1$ *and* $\overline{\pi} + \nu < 1$ *for some* $\nu > 0$*. The productivity of the second type is independently and uniformly distributed on* $[0, 1]$*. Let* N *denote the size of the consideration set,* N_1 *the number of applicants selected from the first type, and* N_2 *the number of the second type selected. Lemma A.1 shows that* $N_1 = 0$ *for small* N *and* $N_1 = 0$ *for large* N*.*

When the consideration set is small, the mean of the parent distribution heavily influences the distribution of the second highest valuation, inducing the committee to select from a distribution with a high mean. When the consideration set is large, the right tail of the parent distribution assumes greater importance, so the committee tends to select from distributions with higher variances. Because the distributions cannot be ranked by FOSD, the sequence of sets $\mathbb{C}^N$ is not monotone increasing.

Example 8 (Divergent preferences) *There are three applicants* a_k *where* $k \in \{1, 2, 3\}$ *and* $\eta = 0$*. The committee's valuations are independently drawn from uniform distributions*

[11] That is, the epigraph for the negative of the mapping formed from joining adjacent points on the graph $(N, L(\mathbb{C}_N))$ with linear segments is not convex.

with support $[\underline{\pi}_k, 1]$ *where* $0 = \underline{\pi}_1 < \underline{\pi}_2 < \underline{\pi}_3$. *The manager, however, values the second applicant at* 1 *and the other two the same way as the committee. Lemma A.2 shows the committee selects* $\mathbb{C}_2 = \{a_1, a_3\}$ *if* $\underline{\pi}_2 - \underline{\pi}_1$ *is small enough.*

Here, depending on the values of $\underline{\pi}_2$ and $\underline{\pi}_3$, rather than jeopodize the chances of the committee's most promising candidate, by exposing the third candidate to competition from the second, whom the manager would select if she could, the committee may select the first and third candidates for the consderation set, instead of its top two. Importantly, both the committee and the manager with the same information would rank applicants by FOSD, but their ordering differs.

4 Identification and Estimation

The portion of the sample panel on vacancies and applications is only five years. Most employees do not switch jobs in the sample frame, let alone more than once. For this reason we impose a heirarchical or ladder structure on how careers evolve within this firm. Because job turnover is a low frequency event in this sample, we assume that conditional on the worker's sociodemographic descriptives, he climbs a career ladder, one rung at a time. As an institutional norm for this firm, this approximation is also a very plausible. Fortunately the matching panel on firm employees recording wages and tenure on the job is longer, providing the means to estimate spell durations. This section parameterizes the model, explains the basis for its identification, and describes elaborates estimation.

Parameterizing the model The unit of observation in the sample is a job vacancy denoted by $n \in \{1, 2, \ldots, N\}$. The applicant pool for the n^{th} vacancy is denoted by $\mathbb{A}_n$. To facilitate identification and estimation we assume that when a new employment opportunity arises, workers do not know how the identity of their rivals. Their information is limited to a partial description of the vacancy, their own personal socio-economic demographics, and their current employment benefits. We represent by (x_{an}, ξ_{an}) the information $a_n \in \mathbb{A}_n$ has at the time a new job opportunity arrives, where x_{an} is a row vector of characteristics relating to the job and the applicant and ξ_{an}, the submission cost, is distributed independently and identically as a logistic random variable. We parameterize the preferences a_n for the job he currently holds by setting:

$$u_{an} = x_{an}\gamma \tag{10}$$

and similarly, his preferences for the job n^{th} opportunity by $\widehat{u}_{an} = \widehat{x}_{an}\gamma$. The relocation cost, $\sigma_\epsilon \epsilon_{an}$, is only revealed to him upon reaching the consideration set. We assume ϵ_{an} is

an independent and identically distributed centered normal random variable.

We denote by $\mathbb{C}_n$ the consideration set for filling the n^{th} job vacancy. The personnel manager evaluates candidates $a_n \in \mathbb{C}_n$ on the basis of (x_{an}, ζ_{an}), where ζ_{an} is a centered normal random variable, independent and identically distributed with standard deviation σ_ξ. We assume:

$$\pi_{an} = x_{an}\alpha + \sigma_\zeta \zeta_{an} \tag{11}$$

where α is a coefficient vector weighting different characteristics skills that a_n brings to the n^{th} position, and that ζ_{an} is a match specific orientation cost the firm incurs by filling job vacancy n with a_n. We assume ζ_{an} is an independent standard normal random variable.

The committee assigned to the n^{th} job vacancy judges candidates $a_n \in \mathbb{B}_n$ on characteristics $(x_{an}, \varepsilon_{an})$, where ε_{an}, is also an independent and identically distributed standard normal random variable, and parameterize the committee's preferences with:

$$v_{an} = x_{an}\beta + \sigma_\varepsilon \varepsilon_{an} \tag{12}$$

Three features differentiate the committee's preferences from the manager's. The committee cares about ε_{an} but the manager does not. The manager cares about ζ_{an} but the committee does not. The magnitudes of σ_ϵ and σ_ζ measure the discrepancy of the differences attributed to these two differences. Third the committee may value x_{an}, differently from the manager, captured by the difference $\beta - \alpha$.

We observe x_{an} for each $a_n \in \mathbb{A}_n$ but $(\xi_{an}, \epsilon_{an}, \zeta_{an}, \varepsilon_{an})$ is an unobserved vector. One of the observed characteristics is whether an applicant is internal, that is currently employed with the firm, or external. We observe the current wage of internal but not external applicants.

Estimation is in two stages. First we estimate the preference parameters of applicants, and the productivity parameters of the firm, an iterative process in which each iteration is simply computed. Then we estimate the committee's preferences from their choice of the consideration set, a computationally intensive estmator that draws upon the estimates obtained from the fist stage.

The Applicants' Preferences and the Firm's Productivity In estimation, the connection between $(\gamma, \sigma_\epsilon)$, the preferences of applicants characterizing the nonpecuniary benefits of the job and the importance of relocation costs, and (α, σ_ζ), the firm's productivity parameters, arises because of two sided selection issues, a many-to-many matching problem. On the supply of this labor market, whether the worker applies for a vacancy brought to his attention partly depends on the relocation costs he would pay conditional on filling the vacancy. The expected relocation cost a would pay for taking the new position, $E[\widehat{\epsilon}_{an}]$, enters

into the applicant submission decision. On the demand side, ζ_{an} is an unobserved productivity parameter benefiting the firm, which partly determines which applicant is selected from the consideration set. Embedded in this microfounded equilibrium is the role of the closest rival for the job, whose relocation cost and productivity parameter helps determines the wage of the hired worker. Consquently these parameters are jointly estimated.

First we estimate some ancillary parameters that relate to thed duration of job spells, the frequency of new employment opportunities, and the probability of being hired conditional on submitting an application and a reservation wage. Then we iterate between solving for the reservation wages of internal candidates given the parameter values for $(\gamma, \sigma_\epsilon, \alpha, \sigma_\zeta)$, and estimating $(\gamma, \sigma_\epsilon, \alpha, \sigma_\zeta)$ given the reservation wages obtained from a previous iteration. Both steps within the iteration have a unique solution, and every fixed point is observationally equivalent. We assume the model is (parametrically) point identified, and find no evidence to the contrary in our computations.

Ancillary parameters and simulations Several nonparametrically identified incidental parameters are estimated in a first stage to be used as input in a second stage. Denote by:

$$p_{an} = \Pr\left[a_n \in \mathbb{B}_n \left| a_n \in \mathbb{A}_n, x_{an}\right.\right]$$

the conditional choice probability of $a_n \in \mathbb{A}_n$ submitting a completed application for the n^{th} vacancy conditional on x_{an} but integrating over ξ_{an}. Also let:

$$\phi_{an} = \Pr\left[\arg\max_{c_n \in \mathbb{C}_n} \left\{M\left(x_{cn}\alpha + \zeta_{cn}, \overline{w}_{cn}\right)\right\} \left| x_{an}, d_{an} = 1\right.\right]$$

denote the probability that a_n would be offered a position to fill the n^{th} vacancy if he applies, given his demographics and his information about the job at the time he is making the decision about whether to apply. Following notation developed in the model let $\boldsymbol{\rho}_{an} = E\left[e^{-\delta\rho_{an}} \left| x_{an}\right.\right]$ and define $\widehat{\tau}_{an} = E\left[e^{-\delta\widehat{\tau}_{an}}\right]$. The parameter vector $(p_{an}, \boldsymbol{\rho}_{an}, \widehat{\tau}_{an}, \phi_{an})$ is ancilliary, identified without imposing any structural assumptions underlying the model, and is estimated in a first step. For expositional convenience we substitute them into the structural estimation equations that we elaborate on below.

The model is estimated in stages, and several of its components exploit simulation. Let $\epsilon_{an}^{(s)}$ denote a normalized relocation shock for the applicant that simulates ϵ_{an}, and denote by $\zeta_{an}^{(s)}$ a productivity shock that simulates ζ_{an}. Similarly, $\varepsilon_{an}^{(s)}$ denotes a simulated committee preference factor standing in for ε_{an}.

Work amenities The nonpecuniary amenities from working in a given position are estimated off the decision by internal applicants to apply for a new job. Lemma 4 expresses the log odds ratio of a_n completing an application for the n^{th} vacancy.

Lemma 9 *Define:*

$$y_{an} = \ln\left(\frac{p_{an}}{1-p_{an}}\right) - \frac{\phi_{an}}{\delta\left[1-\boldsymbol{\rho}_{an}\left(1-\phi_{an}\right)\right]}\left[\begin{array}{l} E\left[\left(1-\boldsymbol{\rho}_{an}\right)\left(\widehat{w}_{an}-w_{an}\right)\right] \\ -\boldsymbol{\rho}_{an}\left\{\sigma_{\epsilon}\widehat{\epsilon}_{an}^{(s)}+\ln\frac{p_{an}}{1-\widehat{p}_{an}}\right\} \end{array}\right] \tag{13}$$

where $E\left[\widehat{\epsilon}_{an}\right] = E\left[\widehat{\epsilon}_{an}^{(s)}\right]$ *. Define* $\nu_{an} = \boldsymbol{\rho}_{an}\sigma_{\epsilon}\left(E\left[\widehat{\epsilon}_{an}\right] - \widehat{\epsilon}_{an}^{(s)}\right)$ *The assumptions of the model imply:*

$$y_{an} = \left(\widehat{x}_{an} - x_{an}\right)\gamma + \nu_{an} \tag{14}$$

If $E\left[\widehat{\epsilon}_{an}\right]$ was known we could identify γ, the nonpecuniary benefits of amenities on the job, off (14) and estimate the parameter vector using ordinary least squares. However $E\left[\widehat{\epsilon}_{an}\right]$ is the expected value of the relocation cost conditional on being offered the new position after submitting a reservation wage, a complication that compels us to estimate γ jointly with α, the coefficients determining the worker's productivity. An iterative procedure was used. To obtain $\gamma^{(k+1)}$ from the k^{th} iteration, write $E^{(k)}\left[\widehat{\epsilon}_{an}\right]$ for the approximation of $E\left[\widehat{\epsilon}_{an}\right]$ after k iterations, form $y_{an}^{(k)}$ by substituting estimates of the ancillary parameters $(p_{an}, \phi_{an}, \boldsymbol{\rho}_{an})$ along with $E^{(k)}\left[\widehat{\epsilon}_{an}\right]$ and a sample analogue of $\widehat{w}_{an}$ into the expression for y_{an} given in (13), and run ordinary least squares.

Productivity To estimate the productivity parameters for the firm, (α, σ_{ζ}), we combine data from both samples described in Section 2. The firm data contains information on internal candidates about their current position and wage, used to estimate their reservation wage. However this sample does not contain the wages of external cadidates for the vacancy, hardly surprising because legal restrictions prevent the firm from using such information. The search data surveys reservation wages. We match an estimated probability distribution characterizing workers searching for a job in the search data with a corresponding distribution for external candidates in the firm data in order to simulate the distribution of reservation wages of external candidates in the firm data.

We exploit two equations in the model to estimate (α, σ_{ζ}). One set of moments is formed by interacting the characteristics of the consideration set with the probability a an internal candidate from the consideration fills the vacancy. A second set of moments is formed from interacting the characteristics of the consideration set with the compensation a winning candidate, is offered. Although the matching algorithm draws together candidates with similar

features, in this way we limit the effect of the search data on our estimates. Specifically, we do not exploit the probability that a particular external candidate is successful; his (simulated) reservation wage appears in estimation only when an internal candidate beats him, as well as the wage equations for both internal and external hires.

Let $\mathbb{C}_n^{(internal)} \subseteq \mathbb{C}_n$ denote the set of internal candidates in the consideration set for the n^{th} vacancy, and let d_{an} denote an indicator whether $a_n \in \mathbb{C}_n^{(internal)}$, an internal applicant, is the successful or not. Write d_n for the vector of indicator variables formed from d_{an} for all $a_n \in \mathbb{C}_n^{(internal)}$. Also let $d_{an}^{(k)}$ indicate whether a simulated $a_n^{(s)} \in \mathbb{C}_n^{(internal)}$ is the successful applicant or not at the k^{th} iteration. That is:

$$d_{an}^{(k)} = 1\left\{\arg\max_{c_n \in \mathbb{C}_n}\left[\left(1-\widehat{\tau}_{cn}\right)\left(x_{cn}\alpha + \sigma_\zeta \zeta_{cn}^{(s)} - \overline{w}_{cn}^{(k)}\right) + \widehat{\tau}_{cn}\delta M_0\right]\right\}$$

and, analogous to the sample counterpart d_n, write $d_n^{(k)}\left(\alpha, \sigma_\zeta, M_0\right)$ for the vector of indicator variables formed from $d_{an}^{(k)}$ for all $a_n^{(s)} \in \mathbb{C}_n^{(internal)}$. From x_n form z_n, a vector of characteristics describing the n^{th} vacancy and the composition of $\mathbb{C}_n$. At the k^{th} iteration we could minimize the Euclidean distance between sample analogues to $E\left[d_n z_n\right]$ and $E\left[d_n\left(\alpha, \sigma_\zeta, M_0\right) z_n\right]$. Instead we augment the moments from the first order condition to this minimization problem with moments obtained from the wage equation.

Since the wages of the hired worker depends on the reservation wage of his closest rival of the job, whose identity is unobserved, we simulate the wage distribution generated by different parameter values for $(\alpha, \sigma_\zeta, M_0)$, and compare moments using the wage data with their corresponding simulations. Given the simulation $\zeta_{cn}^{(s)}$, the parameter values $\left(\alpha^{(k)}, \sigma_\zeta^{(k)}, M_0^{(k)}\right)$ obtained from the k^{th} iteration, along with $\left(\overline{w}_{cn}^{(k)}, d_{cn}^{(k)}\right)$, the simulated reservation wages of candidates in the consideration set, and the induced ranking by the personel manager, we define the closest rival to the successful candidate as:

$$b_n^{(k)} = \arg\max_{c_n \in \mathbb{C}_n}\left\{\left(1-d_{cn}^{(k)}\right)\left[\left(1-\widehat{\tau}_{cn}\right)\left(x_{cn}\alpha^{(k)} + \sigma_\zeta^{(k)}\zeta_{cn}^{(s)} - \overline{w}_{cn}^{(k)}\right) + \widehat{\tau}_{cn}\delta M_0^{(k)}\right]\right\}$$

Let $\overline{w}_n^{(k)}\left(\alpha, \sigma_\zeta\right) = \overline{w}_{bn}^{(k)}$ denote the reservation wage of $b_n^{(k)}$. Define:

$$\tau_{abn}^{(k)} = \left\{1 - E\left[e^{-\delta\overline{\tau}_{bn}^{(k)}}\right]\right\} \Big/ \left\{1 - E\left[e^{-\delta\widehat{\tau}_{an}^{(k)}}\right]\right\}$$

From (**??**) and (11), when $a_n^{(k)}$ is offered the job, his wage is:

$$\begin{aligned}\widehat{w}_{an}^{(k)}\left(\alpha,\sigma_{\zeta},M_0,\sigma_{\epsilon}\right) &= \tau_{abn}^{(k)}\overline{w}_{bn}^{(k)}+\left(x_{an}^{(k)}-\tau_{abn}^{(k)}x_{bn}^{(k)}\right)\alpha-\left(1-\tau_{abn}^{(k)}\right)\delta M_0 \\ &\quad +\sigma_{\varsigma}\zeta_{an}^{(s)}-\tau_{abn}^{(k)}\sigma_{\varsigma}\zeta_{bn}^{(s)}\end{aligned} \tag{15}$$

Given reservation wages for everyone in the consideration set, at the $(k+1)^{th}$ iteration, $\left(\alpha^{(k+1)},M_0^{(k+1)},\sigma_{\varsigma}^{(k+1)},\sigma_{\epsilon}^{(k+1)}\right)$ minimizes:

$$\frac{1}{N}\sum_{n=1}^{N}z_n'\left[\begin{array}{c} d_n-d_n^{(k)}\left(\alpha,\sigma_{\zeta}\right) \\ \widehat{w}_n-\widehat{w}_{an}^{(k)}\left(\alpha,\sigma_{\zeta},M_0,\sigma_{\epsilon}\right)\end{array}\right]'\Omega\frac{1}{N}\sum_{n=1}^{N}z_n'\left[\begin{array}{c} d_n-d_n^{(k)}\left(\alpha,\sigma_{\zeta}\right) \\ \widehat{w}_n-\widehat{w}_{an}^{(k)}\left(\alpha,\sigma_{\zeta},M_0,\sigma_{\epsilon}\right)\end{array}\right]'$$

where Ω is positive semidefinite conforming to $E\left[d_nz_n\right]$. Clearly $E\left[\widehat{w}_{an}z_n\right]=E\left[\widehat{w}_{an}^{(k)}z_n\right]$ and $E\left[d_nz_n\right]=E\left[d_n^{(k)}z_n\right]$. It is straightforward to verify that, given the true reservation wages, the solution to this optimization problem converges to the true parameter values where.

Reservation wages To complete an iteration we update $E^{(k)}\left[\widehat{\epsilon}_{an}\right]$ and $\overline{w}_{cn}^{(k)}$ given $\left(\alpha^{(k)},\sigma_{\zeta}^{(k)},M_0^{(k)},\sigma_{\epsilon}^{(k)}\right)$. The reservation wage depends on the current wage, but in the firm data, which tracks internal applicants, only the latter is observed. Given the parameterization embodied in (10), suppose for the moment that the parameters determining nonpecuniary benefits of jobs are given $\left(\gamma^{(k)},\sigma_{\epsilon}^{(k)}\right)$. Appealing to (3), note that for internal candidates $\left(w_{an},x_{an},\widehat{x}_{an}\right)$ are data, $\boldsymbol{\rho}_{an}$ is an ancillary parameter estimated outside the model, and ϵ_{an} is simulated. In the appendix a lemma on the difference value fucntions derive an expression for the difference in value functions $\left(V_{an}-\overline{V}_{an}\right)$ in terms of the CCPs and other ancillary parameters, along with the coefficent on nonpecuniary benefits, γ. Then (3) implies the reservation wage can be expressed as:

$$\begin{aligned}\overline{w}_{an} &= w_{an}+\left(x_{an}-\widehat{x}_{an}\right)\gamma^{(k+1)} \\ &\quad +\left(1-\boldsymbol{\rho}_{an}\right)^{-1}\delta\left[\boldsymbol{\rho}_{an}\left(V_{an}-\overline{V}_{an}\right)+\sigma_{\epsilon}^{(k+1)}\epsilon_{an}^{(s)}\right]\end{aligned} \tag{16}$$

The Appendix shows the difference $\left(V_{an}-\overline{V}_{an}\right)$ can be expressed as a function of the conditional choice probabilities, which in turn depend on $\overline{w}_{an}$. Clearly $\overline{V}_{an}$ is decreasing in $\overline{w}_{an}$ and hence the solution to (16) is unique. Having computed $\overline{w}_{cn}^{(k)}$, we then compute who is offered the job using $M\left(x_{cn}\alpha+\zeta_{cn},\overline{w}_{cn}\right)$ and hence the winning $\epsilon_{an}^{(s)}$, which gives a simulated draw for $\widehat{\epsilon}_{an}$.

In the search survey data both are observed. For this reason we adopt different procedures for mak Current wages of external applicants are observed. However the reservation wages

of the matched sample are observed.

Committee preferences The committee's preferences are estimated with a simulated methods of moments estimator (SMM). We simulate $\varepsilon_{an}^{(s)}$, the applicant specific disturbance in committee's utility, and draw upon estimates obtained of the applicant and the manager in the previous steps. Given a value for $(\beta, \sigma_\varepsilon, \eta)$, the parameters representing committee preferences, we solve for the optimal consideration sets. Varying $(\beta, \sigma_\varepsilon, \eta)$ we then compare a vector of sample moments of key feaures of the simulated consideration sets with their corresponding sample moments, matching the two vectors by minimizing a quadratic form. The components of the vector comprises

The most challenging part of this component to estimation is computing the committee's optimal consideration set, which is nested inside the estimation algorithm. We do not assume the three conditions used in proving Theorem 2 hold in the applicant population. However our estimator exploits a direct implication of Theorem 2 that within partitioned subsets of the application pool where the preferences of committee and the manager are aligned (for applicants with the same covariates), the marginal approach to adding an extra applicant from within the subset applies. This economizes on how many consideration sets must compared to determine the optimum. The scope for differentiating the preferences of the manager from the committee declines with the size of the consideration set for two reasons. First, fewer candidates the manager might favor are excluded from the set. Second, the committee selects a larger consideration set only when it wants the manager to have a greater say because it cares more about those characteristics it cannot observe in a a similar way to the manager. In the limit they coincide. As a practical compromise, we fully solved the optimal consideration set for up to seven applicants, and then added applicants following the marginal approach of the theorem, checking locally for concavity.

5 Results

In this section, we show the results of the model estimation. First, we show the non-pecuniary preferences of the candidates, which informs the estimation of reservation wages. Next, we show the results of the committee preferences, and last, we will show the results of the productivity parameters, which inform the equilibrium wage determination.[12]

[12]For now, we are only reporting the estimated parameters. We are working to compute bootstrapped standard errors.

5.1 Non-pecuniary parameters

Candidates have preferences over non-pecuniary characteristics. We parameterize this by dividing all jobs into 5 groups based on the division of the position, and assume that the candidate's preferences for a division depend on their demographic characteristics. This allows for women to prefer different types of jobs than men, for example. Table 10 shows the estimated values of the non-pecuniary parameters.

5.2 Committee preferences

Recall that the interview committee has preferences over candidate characteristics that may differ from the productivity terms. Each candidate also has a term that is only observed by the interview committee, η, which we assume to have mean 0 and estimate the variance. The parameters of the preferences of the committee are shown in Table 11.

5.3 Productivity parameters

Table 12 shows the productivity parameters. The first components determine the productivity of a candidate, which is a function of observed characteristics and an unobserved term ζ. We estimate the impact of each characteristic and the standard deviation of ζ. We also estimate the mean and standard deviation of relocation costs ϵ, and the value of a vacancy $M0$.

5.4 Model Fit

Table 13 shows the fit of the model. The first 2 columns show the share of interviewed candidates with a given demographic characteristic, and we see that the model fits the data quite well. We slightly underestimate the number of interviews conducted, as the model predicts an average of 4.2 interviews per position and an average of 5.4 candidates are interviewed for each position in the data. The last 2 columns show the characteristics of the hired candidates, and again we fit the data quite well. In addition, the average wages predicted by the model are quite close to the wages reported in the data.

6 Counterfactuals

In this section, we compute different counterfactuals to understand how hiring decisions would change in alternative environments.

How important is the agency problem? Suppose that the interview committee had the same preferences as the firm when selecting which candidates to interview. The regime change has an immediate impact on who is hired because of its effect on the set of applicants interviewed. This counterfactual reveals the scope of the divergent preferences of the interview committee. It also shows the surplus lost to the firm that occurs because they outsource interview decisions to individuals other than the manager.

To implement this counterfactual, we set the divergent preferences of the interview committee to 0. Table 14 shows how this impacts the demographic characteristics of the chosen consideration set. For most characteristics, there is not a large change. However, we see an almost doubling of the share of black candidates interviewed in this counterfactual. Given that more black candidates are interviewed under this counterfactual, we next ask how it impacts the share of black candidates that are hired, and find that it jumps from 2.2% to 5.6%.

What is the role of multi-stage choice? In our setting, the hired candidate is chosen through a multi-stage process. First, a set of candidates is chosen to be interviewed and after the interviews occur, a candidate is chosen to be hired. We explore alternative mechanisms to understand how the selection mechanism affects who is hired, as well as the surplus of the firm from the hired candidate. We first consider an environment where the interview set is selected randomly, eliminating the committee and manager preferences in the choice of whom to interview. We compare how many interviews are needed to earn the firm the same surplus as in the baseline model. In the baseline, the firm interviews an average of 4.2 candidates. In the counterfactual where consideration sets are chosen randomly, it takes an average of 4.9 interviews to reach the same surplus as in the baseline.

We also consider a mechanism where candidate's are chosen (randomly) sequentially. First, the committee decides whether or not to interview a candidate. If they choose not to interview them, they move on to the next candidate. The manager decides whether or not to offer the position to each interviewed candidate. If they do not offer the position to the interviewed candidate, then the committee moves on to the next candidate. We will consider how this mechanism affects the characteristics and productivity of the hired candidate compared to the baseline mechanism. We are currently working to compute the results of this counterfactual exercise.

What are the effects of regulating hiring practice? Change by fiat, enforced internally by the firm or by law, is a policy tool that will affect hiring outcomes. For example, a quota system might target race and gender groupings both at the interview stage and in

the hiring stage. How would demographic groups be affected if minority and gender equality proportional to the application rates was guaranteed at the interview stage? This approach is not limited to counterfactual policies already debated in public debate; it is also a catalyst for thinking about new ways to regulate hiring. For example, what are the effects of a quota system that even the odds of African Americans receiving as offer, conditional on applying for the job and being minimally qualified, rather than conditional on receiving an interview? Another approach would be a race and gender blind approach to the selection of the interview set, eliminating these characteristics from impacting the probability that a person is interviewed. We simulate this policy to see how it affects who is hired. Table 14 shows this policy has the greatest impact on the share of black candidates interviewed. This policy almost doubles the number of African-American candidates hired (2.2% to 4. 2%).

References

Altonji, J. and Blank, R. (1999). Race and gender in the labor market. *Handbook of Labor Economics*, Volume 30:3143–3259.

Babcock, L. and Laschever, S. (2007). *Women don't ask: The high cost of avoiding negotiation – and positive strategies for change*. Bantam.

Barseghyan, L., Coughlin, M., Molinari, F., and Teitelbaum, J. C. (2021). Heterogenous choice sets and preferences. *Econometrica*, Forthcoming.

Bertrand, M. and Mullainathan, S. (2004). Are emily and greg more employable than lakisha or jamal? a field experiment on labor market discrimination. *The American Economic Review*, 94(4):991–1013.

Blau, F. D. and Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, 55(3):789–865.

Burdett, K. and Mortensen, D. T. (1998). Wage differentials, employer size, and unemployment. *International Economic Review*, 39(2):257–273.

Carneiro, P., Heckman, J. J., and Masterov, D. V. (2005). Labor market discrimination and racial differences in premarket factors. *The Journal of Law and Economics*, 48(1):1–39.

Coughlin, M. (2023). Insurance choice with non-monetary plan attributes: Limited consideration in medicare part d. Working paper.

Faberman, R. J., Mueller, A. I., Sahin, A., and Topa, G. (2022). Job search behavior among the employed and non-employed. *Econometrica*, 90(4):1743–1779.

Gayle, G.-L. and Golan, L. (2012). Estimating a dynamic adverse-selection model: Labour-force experience and the changing gender earnings gap 1968-1997. *The Review of Economic Studies*, 79(1):227–267.

Gayle, G.-L., Golan, L., and Miller, R. A. (2012). Gender differences in executive compensation and job mobility. *Journal of Labor Economics*, 30(4):829–871.

Golan, L., James, J., and Sanders, C. (2024). Estimating a life-cycle generalized roy model of pay and task assignment: Compensating differentials, discrimination and racial gaps. Working paper.

Gottardi, P., Lester, B., Michaels, R., and Wolthoff, R. (2025). Signals, interviews, and discrimination in the labor market. Working paper.

Jovanovic, B. (1979). Job matching and the theory of turnover. *Journal of Political Economy*, 87(5):972–990.

Lazear, E. P. (2000). Performance pay and productivity. *American Economic Review*, 90(5):1346–1361.

Lazear, E. P., Shaw, K. L., and Stanton, C. (2016). Making do with less: Working harder during recessions. *Journal of Labor Economics*, 34(S1, Part 2):S333–S360.

Lazear, E. P., Shaw, K. L., and Stanton, C. T. (2015). The value of bosses. *The Journal of Labor Economics*, 33(4):823–861.

Lentz, R., Piyapromdee, S., and Robin, J.-M. (2023). The anatomy of sorting - evidence from danish data. *Econometrica*, 91(6):2409–2455.

Mortensen, D. T. and Pissarides, C. A. (1994). Job creation and job destruction in the theory of unemployment. *The Review of Economic Studies*, 61(3):397–415.

Neal, D. A. and Johnson, W. R. (1996). The role of premarket factors in black-white wage differences. *Journal of Political Economy*, 104(5):869–895.

O'Neill, D. M. (1970). The effect of discrimination on earnings: Evidence from military test score results. *Journal of Human Resources*, 5(4):475–486.

O'Neill, J. (1990). The role of human capital in earnings differences between black and white men. *Journal of Economic Perspectives*, 4(4):25–45.

Pissarides, C. A. (1985). Short-run equilibrium dynamics of unemployment, vacancies, and real wages. *American Economic Review*, 75(4):676–690.

Postel-Vinay, F. and Robin, J.-M. (2002). Equilibrium wage dispersion with worker and employer heterogeneity. *Econometrica*, 70(6):2295–2350.

Russo, G. and van Ommeren, J. (1998). Gender differences in recruitment outcomes. *Bulletin of Economic Research*, 50(2):0307–3378.

Shimer, R. (2005). The cyclical behavior of equilibrium unemployment and vacancies. *American Economic Review*, 95(1):35–49.

Shukla, S. (2024). Making the elite: Top jobs, disparities, and solutions. Working paper.

Vohra, A. and Yoder, N. (2024). Matching with costly interviews: The benefits of asynchronous offers. Working paper.

Xiao, P. (2021). Wage and employment discrimination by gender in labor market equilibrium. Working paper.

Tables and figures

Table 1: Summary statistics

Summary statistics on applicants	
Share of sample that is African American	4.50%
Share of sample that is female	54.61%
Average number of applications/candidate	3.92
Number of applicants	62,405
Summary statistics on jobs	
Average number of applications per job	49.19
Average number of people interviewed per job	7.38
Number of jobs	5,365
Summary statistics on applications	
Share of applications with no experience at the firm	91.22%
Number of applications	252,256

Table 2: Number of applications submitted

	All	Qualified and interested
	(1)	(2)
African American	0.701***	0.341***
	(0.0715)	(0.0465)
Female	0.108***	0.0922***
	(0.0306)	(0.0199)
Duration working for firm	0.140***	0.131***
	(0.0254)	(0.0165)
Observations	87394	87394

Standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. The dependent variable is the number of applications submitted. Controls for year, education, and age are included but not reported. We additionally control for the square of the duration working for the firm.

Table 3: Average salary of job posting

	All		Qualified and interested	
	(1)	(2)	(3)	(4)
African American	-1072.7***	-523.3	-1242.5***	-1229.1***
	(64.17)	(392.9)	(75.30)	(461.2)
Female	-561.2***	-2031.4***	-613.9***	-2236.5***
	(34.30)	(173.4)	(39.31)	(196.6)
Number of applications submitted per year	-11.76***	-145.2***	-42.10***	-165.6***
	(1.516)	(16.74)	(1.926)	(19.83)
Experience at firm	58.64*	-412.1***	108.5***	-452.4***
	(30.93)	(81.74)	(34.70)	(93.67)
Experience in division	-43.14	123.3	-64.06	202.5*
	(61.50)	(101.7)	(65.96)	(114.1)
Worked for firm in prior year	494.2***		644.2***	
	(122.7)		(136.3)	
Worked for division in prior year	1118.3***	1357.9***	1219.8***	1425.7***
	(210.5)	(305.7)	(226.6)	(342.9)
Previous year salary		0.174***		0.186***
		(0.00478)		(0.00540)
Fulltime position	3657.2***	1780.4***	3315.5***	1366.4**
	(100.2)	(584.0)	(102.9)	(607.9)
Job at central location	-157.9***	1516.2***	-233.4***	1589.2***
	(42.53)	(217.8)	(49.60)	(246.7)
Observations	225138	10420	157197	7976

Standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. The job posting reports the minimum and maximum salaries. The dependent variable is the average of those 2 values. We additionally control for education and age, and include division and occupation fixed effects.

Table 4: Probit regression on being not qualified or not interested in a position

	(1)	(2)
	Not interested	Not qualified
African American	0.0106***	0.0193***
	(0.00268)	(0.00157)
Female	0.00729***	-0.0000539
	(0.00148)	(0.000784)
Number of applications submitted per year	0.000750***	0.000987***
	(0.0000593)	(0.0000456)
Number applications for job	-0.0000815***	-0.000357***
	(0.00000606)	(0.0000132)
Experience at firm	-0.00298**	-0.00276***
	(0.00131)	(0.000641)
Experience in division	-0.00251	-0.00249
	(0.00263)	(0.00159)
Worked for firm in prior year	-0.000416	-0.00214
	(0.00561)	(0.00275)
Worked for division in prior year	-0.0333***	-0.0109**
	(0.00926)	(0.00550)
Fulltime position	0.0407***	0.0286***
	(0.00561)	(0.00310)
Job at central location	0.0484***	0.0129***
	(0.00180)	(0.00104)
Average salary of posted job	-0.00000327***	0.00000278***
	(0.000000106)	(0.000000108)
Observations	224248	217802

Standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. We report marginal effects from a probit regression. We also control for education and age, and include year, division, and occupation fixed effects.

Table 5: Probit regression on being interviewed for a position

	(1)	(2)	(3)	(4)
	All	One hired	Internal	External
African American	-0.0315***	-0.0461***	-0.0531***	-0.0285***
	(0.00505)	(0.00728)	(0.0203)	(0.00494)
Female	-0.00253	0.00639*	-0.00151	-0.00311
	(0.00251)	(0.00351)	(0.00956)	(0.00248)
Number of applications per year	-0.00523***	-0.00460***	-0.00783***	-0.00511***
	(0.000123)	(0.000162)	(0.000563)	(0.000125)
Number of candidates	-0.000681***	-0.00227***	-0.000725***	-0.000460***
	(0.0000153)	(0.0000439)	(0.0000669)	(0.0000126)
Experience at firm	-0.00115	-0.00351	-0.00457***	
	(0.00206)	(0.00262)	(0.00176)	
Experience in division	0.0238***	0.0189***	0.0338***	
	(0.00387)	(0.00535)	(0.00390)	
Worked for firm in prior year	-0.000943	0.00910	-0.00391	
	(0.00859)	(0.0112)	(0.0109)	
Worked for division in prior year	0.102***	0.128***	0.129***	
	(0.0135)	(0.0188)	(0.0167)	
Fulltime position	0.0636***	0.0381***	0.0661**	0.0543***
	(0.00680)	(0.0101)	(0.0260)	(0.00675)
Job at central location	-0.0471***	-0.0607***	-0.0535***	-0.0491***
	(0.00315)	(0.00426)	(0.0120)	(0.00311)
Average salary of posted job	0.00000166***	0.00000145***	0.00000215***	0.00000173***
	(0.000000157)	(0.000000195)	(0.000000540)	(0.000000157)
Observations	158186	87233	14752	149597

Standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. We report marginal effects from a probit regression. We also control for eduation and age, and include year, division, and occupation fixed effects. Number of candidates refers to the number of qualified and interested candidates for a given position.

Table 6: Probit regression on being hired

	(1)	(2)	(3)	(4)
	All	One hired	Internal	External
African American	0.0228**	0.00245	0.0595**	0.0181*
	(0.00965)	(0.0111)	(0.0268)	(0.0102)
Female	0.00834*	0.0170***	0.0204*	0.00592
	(0.00443)	(0.00475)	(0.0118)	(0.00474)
Number of applications per year	-0.00929***	-0.00749***	-0.0141***	-0.00914***
	(0.000353)	(0.000406)	(0.00121)	(0.000367)
Number of people interviewed	-0.00219***	-0.00841***	-0.00250***	-0.00214***
	(0.0000825)	(0.000254)	(0.000302)	(0.0000843)
Experience at firm	0.00246	0.00437	0.00162	
	(0.00355)	(0.00343)	(0.00201)	
Experience in division	-0.0115**	-0.00946*	-0.00594*	
	(0.00531)	(0.00507)	(0.00329)	
Worked for firm in prior year	-0.00695	-0.0144	0.000128	
	(0.0153)	(0.0152)	(0.0158)	
Worked for division in prior year	0.0755***	0.0587***	0.0606***	
	(0.0204)	(0.0197)	(0.0201)	
Job at central location	0.0282***	0.0186***	-0.00310	0.0331***
	(0.00549)	(0.00560)	(0.0143)	(0.00585)
Fulltime position	0.0350***	-0.0134	0.0197	0.0352***
	(0.0128)	(0.0147)	(0.0380)	(0.0135)
Average salary of posted job	5.40e-08	0.000000272	-0.000000551	0.000000103
	(0.000000242)	(0.000000226)	(0.000000578)	(0.000000265)
Observations	38257	22783	4790	33664

Standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. We report marginal effects from a probit regression. We also control for eduation and age, and include year, division, and occupation fixed effects.

Table 7: Durations at job

	(1)	(2)
	OLS	Tobit
African American	-0.929***	-1.189***
	(0.215)	(0.288)
Female	-0.295***	-0.754***
	(0.101)	(0.135)
Year job started		0.810***
		(0.00994)
Observations	11347	11347

Standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. We additionally control for education, unemployment and labor market participation in a given year, and include division and occupation fixed effects.

Table 8: Log earnings regression

	(1)	(2)	(3)	(4)
	All		2007-2011	
African American	-0.100***	-0.137***	-0.110*	-0.110*
	(0.0159)	(0.0315)	(0.0604)	(0.0604)
Female	-0.0926***	-0.0581***	-0.0595***	-0.0598***
	(0.00614)	(0.0125)	(0.0228)	(0.0228)
Duration in current division	0.0960***	0.209***	0.118***	0.118***
	(0.00172)	(0.00404)	(0.00886)	(0.00887)
Experience squared	-0.00339***	-0.00900***	-0.00421***	-0.00420***
	(0.0000766)	(0.000240)	(0.000536)	(0.000536)
Switches divisions	0.144***		0.0818	0.0796
	(0.0161)		(0.0631)	(0.0632)
New job		0.216***		
		(0.0172)		
Number applications submitted			-0.00484*	
			(0.00265)	
Number qualified/interested submitted				-0.00552
				(0.00376)
Constant	9.772***	9.176***	9.511***	9.510***
	(0.0279)	(0.0409)	(0.0755)	(0.0756)
Observations	50524	15790	3153	3153

Standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. We also include division and year fixed effects. Controls for education and age are included but not reported. Experience is calculated using the number of years a person is observed continuously in that division.

Table 9: Log earnings regression for job switchers

	(1)	(2)	(3)
African American	-0.0929	-0.0651	-0.0366*
	(0.0719)	(0.0665)	(0.0202)
Female	-0.109***	-0.0896***	-0.0569***
	(0.0335)	(0.0311)	(0.00805)
Duration in current division	0.0729***	0.0539***	0.0414***
	(0.0116)	(0.0108)	(0.00245)
Experience squared	-0.00245***	-0.00183***	-0.00146***
	(0.000716)	(0.000664)	(0.000103)
Switches divisions	0.0852*	0.0690	0.0296
	(0.0470)	(0.0436)	(0.0240)
Number applications for job	-0.000629***	-0.000291**	
	(0.000122)	(0.000115)	
Salary of job posting		0.0000205***	0.0000204***
		(0.00000154)	(0.000000566)
Years since new job			0.0488***
			(0.00542)
Salary of job posting x years since new job			-0.000000356***
			(9.48e-08)
Constant	9.987***	9.132***	9.503***
	(0.184)	(0.182)	(0.0483)
Observations	1212	1208	15384

Standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. We also include division and year fixed effects. Controls for education and age are included but not reported. Experience is calculated using the number of years a person is observed continuously in that division.

Table 10: Candidate non-pecuniary preferences

		Division			
		1	2	3	4
Constant		6306.5	22684.9	-2034.0	-4186.7
Some college	382.8	4242.6	-5911.1	-3351.0	658.2
College	-1162.2	2359.4	-6870.9	920.7	3157.4
Post college	-690.7	8909.3	-19285.6	2280.1	-1972.0
Female	-2786.0	-3910.1	-2363.5	3957.3	4009.5
Black	-2408.5	-20879.3	-8954.2	6003.0	-5184.2

Table 11: Committee preferences

	Estimate
High school	8103.0
Some college	9970.5
College	4232.8
Female	9862.3
Black	-7984.0
Internal	3308.9
Cost of interview	3811.5
Standard deviation of η	9652.2

Table 12: Productivity parameters

	Estimate
High school	-7695.7
Some college	-4212.1
College	11370.1
Female	8613.3
Black	-13,006.9
Internal	-9888.3
Value of a vacancy	9490.1
Mean of ϵ	-6257.8
Standard deviation of ϵ	7700.6
Standard deviation of ζ	110.8

Table 13: Model Fit

	Share of candidates			
	Interviewed		Hired	
	Model	Data	Model	Data
Women	0.55	0.53	0.57	0.57
Black	0.027	0.031	0.031	0.029
Internal to firm	0.071	0.079	0.07	0.11
High school	0.15	0.16	0.14	0.14
Some college	0.32	0.33	0.34	0.32
College	0.40	0.37	0.39	0.39
Post college	0.14	0.13	0.13	0.14
Total number of interviews	4.2	5.4		
Wage of hired candidate			39,135.0	38,306.0

Table 14: Counterfactuals

	Share of interviewed candidates		
	Baseline	No divergent preferences	Race-gender blind interview choice
Women	0.55	0.54	0.53
Black	0.027	0.049	0.047
Internal to firm	0.07	0.058	0.06
High school	0.15	0.16	0.15
Some college	0.32	0.35	0.31
College	0.40	0.35	0.39
Post college	0.14	0.14	0.14

Appendix A: Comparison with ACS/PSID data

Table 15 compares the demographics of our sample to the ACS, using only labor market participants located in the a similar geographic region.

Table 15: Comparison of our sample and the ACS

	Sample	ACS
Age under 18	0.0020	0.021
Age 18-29	0.31	0.21
Age 30-39	0.26	0.19
Age 40-49	0.24	0.24
Age 50-59	0.17	0.23
Age 60-69	0.023	0.089
Age above 70	0.0013	0.022
High school or less	0.22	0.47
Some college	0.38	0.25
College degree	0.32	0.18
Graduate degree	0.091	0.09
African American	0.045	0.063
Female	0.54	0.48
Number of obsservations	62,405	1,727,612

Next, we run an earnings regression using our data and the ACS and the PSID to compare the determinants of wage outcomes. These results are shown in Table 16. All of these regressions include occupation fixed effects, with the same excluded occupation. They also include controls for race, gender, age, education, and year fixed effects. In column (4), we only include individuals in the PSID who work at large (larger than the median) establishments. We also do this using a least absolute deviation regression, those results are in Table 17.

Table 16: OLS wage regressions in our data, the ACS and the PSID

	(1) Our data	(2) ACS	(3) PSID	(4) PSID, large establishments
African American	-3636.7**	-1960.3***	-3661.8***	-4150.7***
	(1768.8)	(124.2)	(1215.7)	(1457.3)
Female	-3113.5***	-13953.0***	-13809.5***	-11037.4***
	(776.7)	(64.53)	(990)	(1281.9)
Age	378.7***	392.0***	452.2***	497.0***
	(30.8)	(2.367)	(35.69)	(47.43)
Post-secondary education	2574.5***	4640.6***	3674.9***	4932.6***
	(964.9)	(72.34)	(1045)	(1374.2)
College degree	4984.1***	17389.3***	20072.3***	19580.8***
	(1058.2)	(82.89)	(1298.7)	(1723.8)
Graduate degree	10746.6***	33656.2***	38025.7***	35745.0***
	-1380.3	(106.2)	(1805.4)	(2403.8)
Constant	22797.8***	34401.9***	34527.0***	32994.6***
	(1938)	(170.5)	(4714.1)	(5203.6)

Occupation and year fixed effects included but not reported. We only include observations with income above a certain threshold to eliminate part time salaries.

Table 17: Least absolute deviations wage regressions in our data, the ACS and the PSID

	(1) Our data	(2) ACS	(3) PSID	(4) PSID, large establishments
African American	-4351.5**	-1126.4***	-2700.0**	-4018.2***
	(2166.2)	(112.3)	(1062.5)	(1288.3)
Female	-3183.1***	-10213.2***	-10100***	-8872.7***
	(951.2)	(58.32)	(865.2)	(1133.2)
Age	261.8***	317.0***	400.0***	454.5***
	(37.73)	(2.14)	(31.19)	(41.93)
Post-secondary education	1746.9	4300.0***	3500.0***	4745.5***
	(1181.7)	(65.38)	(913.3)	(1214.8)
College degree	3059.2**	14064.2***	15300***	15945.5***
	(1296)	(74.92)	(1135)	(1523.8)
Graduate degree	6498.7***	27945.3***	28000.0***	29041.4***
	(1690.4)	(95.96)	(1577.8)	(2124.9)
Constant	25759.4***	29505.7***	37600.0***	31763.6***
	(2373.4)	(154.1)	(4119.9)	(4600)

Occupation and year fixed effects included but not reported. We only include observations with income above a certain threshold to eliminate part time salaries.

Appendix B

Theorem 10 *Assume $L_c = 0$ for all $c \in \mathbb{B}$ and $F_{\mathbb{C}}(\pi_{\mathbb{C}}, \overline{w}_{\mathbb{C}}) = \prod_{c\in\mathbb{C}} F_c(\pi_c, \overline{w}_c)$. Also assume there exists an ordering within $\mathbb{B}$ denoted by $\{c_i\}_{i=1}^{B}$ such that for all $i < B$:*

$$G_i(M) \equiv \int\int_{M(\pi_c,\overline{w}_c)\leq M} dF_c(\pi_c, \overline{w}_c) \leq G_i(M)$$

Denote by $\mathbb{C}^o$ the optimal consideration set and its cardinality by $C^o \equiv |\mathbb{C}^o|$. Then $\mathbb{C}^o = \{c_i\}_{i=1}^{C^o}$ where C^o uniquely solves $\overline{L}(\mathbb{C}^o) \geq \max\{L(\mathbb{C}^o - c_C) + \eta, L(\mathbb{C}^o + c_{C+1}) - \eta\}$.

Proof. There are three parts to the proof. In the first two parts we show $\mathbb{C}^o = \{c_i\}_{i=1}^{C^o}$ for some $C^o \in \{1, \ldots, B\}$, separately considering the cases when $|\mathbb{C}| = 2$ and $|\mathbb{C}| > 2$. Then we solve for C^o. Regarding the first two parts, suppose $G_i(M) \leq G_j(M)$ and $c_j \in \mathbb{C}$ but $c_i \notin \mathbb{C}$.

1. For $|\mathbb{C}| = 2$, write $\mathbb{C} = \{c_h, c_j\}$, let M_h denote the value of c_h, and define $\overline{M}_h \equiv G_j^{-1}[G_i(M_h)]$. For $k \in \{i, j\}$ define the expected value of the committee conditional on M_h as:

$$L(\{c_h, c_k\}|M_h) \equiv \int_0^{M_h} M dG_k(M) + M_h[1 - G_k(M_h)]$$

   Then:

$$L(\{c_h, c_i\}|M_h) - L(\{c_h, c_j\}|M_h) = \int_{\overline{M}_h}^{M_h} (M_h - M) dG_k(M) \geq 0$$

   Integrating over M_h then proves including every $c_i \in \mathbb{C}^o$ FOSD every $c_j \notin \mathbb{C}^o$ for $|\mathbb{C}| = 2$.

2. For $|\mathbb{C}| > 2$ condition on $\mathbb{C}-c_j$, the values of the other elements in the consideration set, letting $M^{(1)}$ and $M^{(2)}$ denote the values of the maximum and second highest values in set $\mathbb{C}-c_j$. For $k \in \{i, j\}$ define the expected value of the committee conditional on $M^{(1)}$ and $M^{(2)}$ as:

$$L\left(\mathbb{C}-c_j+c_k \middle| M^{(1)}, M^{(2)}\right) \equiv M^{(2)} G_k\left(M^{(2)}\right) + \int_{M^{(2)}}^{M^{(1)}} M dG_k(M) + M^{(1)}[1 - G_k(M)]$$

   Define $\overline{M}^{(1)} \equiv G_j^{-1}\left[G_i\left(M^{(1)}\right)\right]$ and similarly define $\overline{M}^{(2)} \equiv G_j^{-1}\left[G_i\left(M^{(2)}\right)\right]$. Also define the mapping $\delta(M)$ to solve $G_i(M + \delta(M)) = G_j(M)$, which is positive for all M. The ordering of M_i and M_j plus FOSD imply:

$$\overline{M}^{(2)} \leq \min\left\{\overline{M}^{(1)}, M^{(2)}\right\} \leq \max\left\{\overline{M}^{(1)}, M^{(2)}\right\} \leq M^{(1)}$$

   There are two possibilities; either $\overline{M}^{(1)} \leq M^{(2)}$ or $M^{(2)} < \overline{M}^{(1)}$.When $M^{(2)} < \overline{M}^{(1)}$ and using the equality $G_i\left(M^{(2)}\right) = G_j\left(\overline{M}^{(2)}\right)$:

$$\begin{aligned} & L\left(\mathbb{C}-c_j+c_i \middle| M^{(1)}, M^{(2)}\right) - L\left(\mathbb{C} \middle| M^{(1)}, M^{(2)}\right) \\ = & \int_{\overline{M}^{(2)}}^{M^{(2)}} \left[M - M^{(2)}\right] dG_i(M) + \int_{\overline{M}^{(2)}}^{\overline{M}^{(1)}} \delta(M) dG_j(M) + \int_{\overline{M}^{(1)}}^{M^{(1)}} \left(M^{(1)} - M\right) dG_j(M) \end{aligned}$$

When $\overline{M}^{(1)} \leq M^{(2)}$ and using the equality $G_i\left(M^{(1)}\right) = G_j\left(\overline{M}^{(1)}\right)$:

$$\begin{aligned} & L\left(\mathbb{C}-c_j+c_i\left|M^{(1)}, M^{(2)}\right.\right) - L\left(\mathbb{C}\left|M^{(1)}, M^{(2)}\right.\right) \\ = & \int_{M^{(2)}}^{M^{(1)}} \left(M - M^{(2)}\right) dG_i(M) + \int_{\overline{M}^{(1)}}^{M^{(2)}} \left[M^{(1)} - M^{(2)}\right] dG_j(M) + \int_{M^{(2)}}^{M^{(1)}} \left(M^{(1)} - M\right) dG_j(M) \end{aligned}$$

In both cases $L\left(\mathbb{C}-c_j+c_i\left|M^{(1)}, M^{(2)}\right.\right) \geq L\left(\mathbb{C}\left|M^{(1)}, M^{(2)}\right.\right)$. Integrating over $M^{(1)}$ and $M^{(2)}$ then proves including every $c_i \in \mathbb{C}^o$ FOSD every $c_j \notin \mathbb{C}^o$ for $|\mathbb{C}| > 2$. This proves $\mathbb{C}^o = \{c_i\}_{i=1}^{C^o}$ for some optimally chosen $C^o \in \{1, 2, \ldots, B\}$.

3. The marginal condition (9) is necessary because otherwise it would be optimal to either shrink or enlarge the consideration set by at least one applicant. Noting the marginal cost of an interview is η, a constant, we establish sufficiency by showing the expected gross benefit of adding an extra applicant to the consideration set declines with its size.[13] That is we now show:

$$E\left[\max\{M_i\}_{i=1}^{I+1} - \max\{M_i\}_{i=1}^{I}\right] \geq E\left[\max\{M_i\}_{i=1}^{I+2} - \max\{M_i\}_{i=1}^{I+1}\right]$$

Defining $M^{(I)} = \max\{M_i\}_{i=1}^{I}$:

$$\begin{aligned} E\left[\max\{M_i\}_{i=1}^{I+1} - \max\{M_i\}_{i=1}^{I}\right] &= E\left[\max\left\{M^{(I)}, M_{I+1}\right\} - M^{(I)}\right] \\ &= E\left[\max\left\{0, M_{I+1} - M^{(I)}\right\}\right] \end{aligned}$$

Also define the random variable:

$$Z_i = \begin{cases} 0 \text{ if } M_{I+i} \leq M^{(I)} \\ M_{I+i} \text{ if } M_{I+i} > M^{(I)} \end{cases}$$

for $i \in \{1, 2\}$. Since $M^{(I)}$ is increasing in I (used in the first inequality below), and Z_1 FOSD Z_2 (used in the second) because $G_{I+1}(\cdot) \leq G_{I+2}(\cdot)$, it now follows that:

$$\begin{aligned} & E\left[\max\{M_i\}_{i=1}^{I+1} - \max\{M_i\}_{i=1}^{I}\right] - E\left[\max\{M_i\}_{i=1}^{I+2} - \max\{M_i\}_{i=1}^{I+1}\right] \\ = & E\left[\max\left\{0, M_{I+1} - M^{(I)}\right\}\right] - E\left[\max\left\{0, M_{I+2} - M^{(I+1)}\right\}\right] \\ \geq & E\left[\max\left\{0, M_{I+1} - M^{(I)}\right\}\right] - E\left[\max\left\{0, M_{I+2} - M^{(I)}\right\}\right] \\ = & \int_{-\infty}^{\infty}\int_{M^{(I)}}^{\infty} M\left[dG_{M_{I+1}}(M) - dG_{M_{I+2}}(M)\right] \\ \geq & 0 \end{aligned}$$

4. See also https://math.stackexchange.com/questions/2439394/are-expected-order-statistics-always-concave-in-sample-size.

■

[13] More precisely, the epigraph of $l(N, \lambda) : \mathbb{R} \times [0, 1] \to \mathbb{R}$ defined as $l(N) = (\lambda - 1) L\left(\mathbb{C}^{N+1}\right) - L\left(\mathbb{C}^N\right)$ is convex.

Lemma 11 *The productivity of the first of two types of applicant is an independent random variable uniformly distributed on* $[\underline{\pi},\overline{\pi}]$ *where* $\underline{\pi}+\overline{\pi}>1$ *and* $\overline{\pi}+\nu<1$ *for some* $\nu>0$. *The productivity of the second type is independently and uniformly distributed on* $[0,1]$. *We denote the composition of the consideration set by* $\mathbb{C}=\{N_1,N_2\}$ *where* N_i *is the number of applicants selected from type* $i\in\{1,2\}$. *Then* $\mathbb{C}^N=\{N,0\}$ *for sufficently mall* N *and* $\mathbb{C}^N=\{0,N\}$ *for sufficiently large* N.

Proof. We first consider the specialization $\underline{\pi}=\overline{\pi}$, implying that every applicant of the first type is identical. Since $\overline{\pi}>1/2$ it immediately follows that $\mathbb{C}^2=\{2,0\}$. There is no gain from setting $N_1>2$. Denote by $L(N_1,N_2)$ the expected value to the committee from selecting the composition (N_1,N_2). Then:

$$L(N_1,N_2)=\begin{cases}\int_0^1 NM^N dM=N/(N+1) & \text{if } N_1=0\\ \Pr\{\text{all type 2}<\overline{\pi}\}\int_0^{\overline{\pi}}\text{highest}\\ \quad+\Pr\{\text{just 1 type 2}>\overline{\pi}\}\,\overline{\pi} & \text{if } N_1=1\\ \quad+\int_{\pi}^1\text{second highest}\\ \Pr\{\text{at most 1 type 2}>\overline{\pi}\}\,\overline{\pi} & \text{if } N_1=2\\ \quad+\int_{\overline{\pi}}^1\end{cases}$$

We show that $L(N,0)>\max\{L(N-1,1),L(N-2,2)\}$ for $N>\overline{N}$. To complete the proof we note that a distribution on support $[\underline{\pi},\overline{\pi}]$ is FOSD dominated by concentrating all the mass at $\overline{\pi}$. Therefore if the second type is drawn from support $[\underline{\pi},\overline{\pi}]$ then $\mathbb{C}^N=\{N,0\}$ for $N>\overline{N}$. ■

Lemma 12 *There are three applicants* $k\in\{1,2,3\}$. *From the committee's perspective, the productivity of each is an independent and uniformly distributed on* $[\underline{\pi}_k,1]$ *where* $0=\underline{\pi}_1<\underline{\pi}_2<\underline{\pi}_3$. *The manager values the first and third applicant the same way as the committee but values the second applicant more with weight* $(1-\alpha)^{-1}$, *where* $1<\alpha<\underline{\pi}_3/\underline{\pi}_2$. *Lemma A2 shows that:*

$$\mathbb{C}^N=\begin{cases}\{0,1,1\} & \textit{if } N=2 \textit{ and } \alpha.\\ \{1,0,1\} & \textit{if } N=2 \textit{ and } \alpha.\\ \{1,1,1\} & \textit{if } N=3 \textit{ and } \alpha.\\ \{1,0,1\} & \textit{if } N=3 \textit{ and } \alpha.\end{cases}$$

Lemma on reservation wage and proof In this application the conditional value functions V_0 and $\overline{V}_1$ are defined

$$\begin{aligned}V_0 &= \delta^{-1}\left[w+u+\rho\left(\delta V-w-u\right)\right]\\ \overline{V}_1-\epsilon &= \delta^{-1}\left[\overline{w}+\widehat{u}+\rho\left(\delta\overline{V}-\overline{w}-\widehat{u}\right)\right]\end{aligned}$$

By definition the reservation wage equates $\overline{V}_1-V_0$ with ϵ. Substituting the expressions for V_0 and $\overline{V}_1$ and multiplying both sides with δ we obtain:

$$\begin{aligned}\overline{w}+\widehat{u}+\rho\left(\delta\overline{V}-\overline{w}-\widehat{u}\right)-\delta\epsilon &= w+u+\rho\left(\delta V-w-u\right)\\ (1-\rho)\left(\overline{w}+\widehat{u}-w-u\right) &= \rho\delta\left(V-\overline{V}\right)+\delta\epsilon\end{aligned}$$

Note that $\overline{V}$ is monotone declining in $\overline{w}$ and hence the solution is unique. The lemma now follows by by dividing though by$(1-\rho)$ and making $\overline{w}$ the subject of the equation.

Lemma on equilibrium wage

$$\widehat{w}_a = \pi_a + \frac{1 - E\left[e^{-\delta\overline{\tau}_b}\right]}{1 - E\left[e^{-\delta\widehat{\tau}_a}\right]}\left(\overline{w}_b - \pi_b\right) + \frac{E\left[e^{-\delta\widehat{\tau}_a}\right] - E\left[e^{-\delta\overline{\tau}_b}\right]}{1 - E\left[e^{-\delta\widehat{\tau}_a}\right]}\delta M_0$$

Proof of lemma on equilibrium wage Suppose $a \in \mathbb{C}$ is offered the job, $b \in \mathbb{C}$ is the next most attractive candidate, and $\overline{w}_b$ is the reservation wage of b. Abbreviating the notation, define $\tau_a = E\left[e^{-\delta\widehat{\tau}_a}\right]$ and $\tau_b = E\left[e^{-\delta\overline{\tau}_b}\right]$, which we interpret as the expected discount factors on an income unit from postponing it until a and b would quit the job they are currently competing for, given wages of $\widehat{w}_a$ and $\overline{w}_b$ respectively. To derive the equilibrium offer to a, which we denote by $\widehat{w}_a$, we equate $M\left(\pi_a, w_a\right)$ with $M\left(\pi_b, \overline{w}_b\right)$. Appealing to (4):

$$\delta^{-1}\left(1 - \tau_a\right)\left(\pi_a - \widehat{w}_a\right) + \tau_a M_0 = \delta^{-1}\left(1 - \tau_b\right)\left(\pi_b - \overline{w}_b\right) + \tau_b M_0$$

Solving for $\widehat{w}_a$ gives:

$$\widehat{w}_a = \pi_a + \frac{1 - \tau_b}{1 - \tau_a}\left(\overline{w}_b - \pi_b\right) + \frac{\tau_a - \tau_b}{1 - \tau_a}\delta M_0$$

$$\begin{aligned}
\widehat{w}_a &= \pi_a + \frac{1 - \overline{\tau}_b}{1 - \widehat{\tau}_a}\left(\overline{w}_b - \pi_b\right) + \frac{\widehat{\tau}_a - \overline{\tau}_b}{1 - \widehat{\tau}_a}\delta M_0 && (17)\\
&= \pi_a + \frac{1 - \widehat{\tau}_a}{1 - \widehat{\tau}_a}\left(\overline{w}_b - \pi_b\right) + \frac{\widehat{\tau}_a - \overline{\tau}_b}{1 - \widehat{\tau}_a}\left(\overline{w}_b - \pi_b\right) + \frac{\widehat{\tau}_a - \overline{\tau}_b}{1 - \widehat{\tau}_a}\delta M_0 && (18)\\
&= \overline{w}_b + \pi_a - \pi_b + \frac{\widehat{\tau}_a - \overline{\tau}_b}{1 - \widehat{\tau}_a}\left(\overline{w}_b - \pi_b + \delta M_0\right) && (19)
\end{aligned}$$

Substituting for τ_a and τ_b gives the result.

Lemma on probability of submission

$$\begin{array}{l}
\ln\left(\frac{p_{an}}{1-p_{an}}\right) \\
-\frac{\phi_{an}}{\delta[1-\boldsymbol{\rho}_{an}(1-\phi_{an})]}
\end{array}
\left[\begin{array}{l}
E\left[\left(1-\boldsymbol{\rho}_{an}\right)\left(\widehat{w}_{an}-w_{an}\right)\right] \\
-\boldsymbol{\rho}_{an}\left\{E\left[\widehat{\epsilon}_{an}\right]+\ln\frac{p_{an}}{1-\widehat{p}_{an}}\right\}
\end{array}\right]
= \left(\widehat{x}_{an} - x_{an}\right)\gamma$$

which comes form

$$\ln\left(\frac{p_{an}}{1-p_{an}}\right) = \frac{\phi_{an}}{\delta\left[1-\boldsymbol{\rho}_{an}\left(1-\phi_{an}\right)\right]}
\left[\begin{array}{l}
E\left[\left(1-\boldsymbol{\rho}_{an}\right)\left(\widehat{u}_{an}+\widehat{w}_{an}\right)\right] \\
-E\left[\left(1-\boldsymbol{\rho}_{an}\right)\left(u_{an}+w_{an}\right)\right] \\
-\boldsymbol{\rho}_{an}\left\{E\left[\widehat{\epsilon}_{an}\right]+\ln\frac{p_{an}}{1-\widehat{p}_{an}}\right\}
\end{array}\right]$$

Proof of lemma on probability of submission To simplify the exposition of the proof we abbreviate the notation by defining:

- $\psi_k = V - V_k$ and $\widehat{\psi}_k = \widehat{V} - \widehat{V}_k$ for $k \in \{0, 1\}$
- $\rho = E\left[e^{-\delta\rho}\right]$ denote the expected discount factor on current values received at next employment opportunity

- $\mu = \delta^{-1} E\left[\left(1 - e^{-\delta\rho}\right)(u + w)\right]$ and $\widehat{\mu} = \delta^{-1} E\left[\left(1 - e^{-\delta\rho}\right)(\widehat{u} + \widehat{w})\right]$
- $\widehat{\epsilon}_{\mathbf{a}} = E\left[\widehat{\epsilon}\right]$

Also let $q(\cdot)$ denote the inverse of the cumulative distribution function for ξ and p the probability that $\xi \leq V_1 - V_0$, implying $q(p) = V_1 - V_0$. Appealing to (**??**) and (**??**) the notation defined above implies:

$$\begin{aligned} V_1 &= (1-\phi)(\mu + \boldsymbol{\rho} V) + \phi\left(\widehat{\mu} + \widehat{\epsilon} + \boldsymbol{\rho}\widehat{V}\right) \\ V_0 &= \mu + \boldsymbol{\rho} V \end{aligned}$$

Differencing:

$$\begin{aligned} V_1 - V_0 &= (1-\phi)(\mu + \boldsymbol{\rho} V) + \phi\left(\widehat{\mu} + \widehat{\epsilon} + \boldsymbol{\rho}\widehat{V}\right) - \mu - \boldsymbol{\rho} V \\ &= \phi\left[\widehat{\mu} - \mu + \widehat{\epsilon} + \boldsymbol{\rho}\left(\widehat{V} - V\right)\right] \end{aligned} \tag{20}$$

But:

$$\begin{aligned} \widehat{V} - V &= \widehat{V}_0 + \widehat{\psi}_0 - V_0 - \psi_0 \\ &= \widehat{\mu} + \boldsymbol{\rho}\widehat{V} + \widehat{\psi}_0 - (\mu + \boldsymbol{\rho} V) - \psi_0 \\ &= \rho\left(\widehat{V} - V\right) + \widehat{\mu} - \mu + \widehat{\psi}_0 - \psi_0 \\ &= \frac{\widehat{\mu} - \mu + \widehat{\psi}_0 - \psi_0}{1 - \boldsymbol{\rho}} \end{aligned} \tag{21}$$

Substituting (21) into (20) yields:

$$\begin{aligned} V_1 - V_0 &= \phi\left[\widehat{\mu} - \mu + \widehat{\epsilon} + \boldsymbol{\rho}\frac{\widehat{\mu} - \mu + \widehat{\psi}_0 - \psi_0}{1 - \boldsymbol{\rho}}\right] \\ &= \frac{\phi}{1 - \boldsymbol{\rho}}\left[(1 - \boldsymbol{\rho})(\widehat{\mu} - \mu + \widehat{\epsilon}) + \boldsymbol{\rho}\left(\widehat{\mu} - \mu + \widehat{\psi}_0 - \psi_0\right)\right] \\ &= \frac{\phi}{1 - \rho}\left[\widehat{\mu} - \mu + (1 - \boldsymbol{\rho})\widehat{\epsilon} - \boldsymbol{\rho}\left(\widehat{\psi}_0 - \psi_0\right)\right] \end{aligned}$$

Since $q(p) = V_1 - V_0$ it now follows that:

$$q(p) = \frac{\phi}{1 - \boldsymbol{\rho}}\left[\widehat{\mu} - \mu + (1 - \boldsymbol{\rho})\widehat{\epsilon} - \boldsymbol{\rho}\left(\widehat{\psi}_0 - \psi_0\right)\right] \tag{22}$$

Under the parameterization $q(p) = \ln p - \ln(1-p)$ while $\widehat{\psi}_0 - \psi_1 = \ln p - \ln \widehat{p}$. (See Arcidiacono and Miller, 2011.) Substituting these expressions along with those for $\widehat{\mu}$ and μ into (22) completes the proof.

Lemma "Difference in Value Functions"

$$V_{an} - \overline{V}_{an} = \left(\overline{V}_{an} - \widehat{V}_{an}\right) + \left(\widehat{V}_{an} - V_{an}\right)$$

where:

$$\overline{V}_{an} - \widehat{V}_{an} = (x_{an} - \widehat{x}_{an})\,\gamma + w - E\left[\widehat{w}\right] + \ln\left(1 - \widehat{p}_{an}\right)/(1 - \overline{p}_{an})$$

and:

$$\widehat{V} - V = \frac{(1-\boldsymbol{\rho})(1-\phi)}{1-\boldsymbol{\rho}(1-\phi)}\delta^{-1}E\left[(\widehat{u}+\widehat{w})\right] - \frac{(1-\boldsymbol{\rho})(1-\phi)}{1-\boldsymbol{\rho}(1-\phi)}\delta^{-1}(u+w) - \frac{\phi\widehat{\epsilon} - \widehat{\psi}_0 + \psi_1}{1-\boldsymbol{\rho}(1-\phi)}$$

Proof of lemma "Difference in Value Functions" Following the previous proofs we drop the subscripts to simplifiy the notation. Following the previous notation V_0 denotes the conditional value function for remaining on the current job. Denote by $\overline{V}_1$, the conditional value function for starting anew, after allowing for the relocation cost of ϵ. Also let $\overline{V}$ denote the social surplus function for a person who was previously hired at their reservation wage. The definition of conditional value functions imply $V_0 = \mu + \boldsymbol{\rho}V$ and $\overline{V}_1 = \overline{\mu} + \boldsymbol{\rho}\overline{V}$. At the reservation wage $V_0 = \overline{V}_1 + \epsilon$. Substituting the expressions for the conditional value functions into this equation:

$$\begin{aligned}
\overline{\mu} &= \mu - \epsilon + \boldsymbol{\rho}\left(V - \overline{V}\right) \\
\overline{\mu} &= \mu - \epsilon + \boldsymbol{\rho}\left(\overline{V} - V\right)
\end{aligned} \tag{23}$$

We seek an expression for:

$$\overline{V} - V = \left(\overline{V} - \widehat{V}\right) + \left(\widehat{V} - V\right)$$

We obtain expressions for the components using the following notation. Let:

$$\begin{aligned}
\overline{\mu} &= \delta^{-1}E\left[(1-\boldsymbol{\rho})(\widehat{u}+\overline{w})\right] \\
\widehat{\mu} &= \delta^{-1}E\left[(1-\boldsymbol{\rho})(\widehat{u}+\widehat{w})\right] \\
\mu &= \delta^{-1}\left\{1 - E\left[e^{-\delta\overline{\rho}}\right]\right\}(w+u)
\end{aligned}$$

But:

$$\begin{aligned}
\overline{V} - \widehat{V} &= \overline{V}_0 + \overline{\psi}_0 - \widehat{V}_0 - \widehat{\psi}_0 \\
&= \overline{V}_0 - \widehat{V}_0 + \overline{\psi}_0 - \widehat{\psi}_0 \\
&= \overline{\mu} + \boldsymbol{\rho}\overline{V} - \widehat{\mu} - \boldsymbol{\rho}\widehat{V} + \overline{\psi}_0 - \widehat{\psi}_0 \\
&= \boldsymbol{\rho}\left(\overline{V} - \widehat{V}\right) + \overline{\mu} - \widehat{\mu} + \overline{\psi}_0 - \widehat{\psi}_0 \\
&= \frac{\overline{\mu} - \widehat{\mu} + \overline{\psi}_0 - \widehat{\psi}_0}{1-\boldsymbol{\rho}}
\end{aligned} \tag{24}$$

Also note:

$$\begin{aligned}
\widehat{V}-V &= \widehat{V}_0+\widehat{\psi}_0-V_1-\psi_1 \\
&= \widehat{\mu}+\boldsymbol{\rho}\widehat{V}+\widehat{\psi}_0 \\
&\quad -(1-\phi)(\mu+\boldsymbol{\rho}V)-\phi\left(\widehat{\mu}+\widehat{\epsilon}+\boldsymbol{\rho}\widehat{V}\right)-\psi_1 \\
&= \rho(1-\phi)\left(\widehat{V}-V\right)+(1-\phi)(\widehat{\mu}-\mu)-\phi\widehat{\epsilon}+\widehat{\psi}_0-\psi_1 \\
&= \frac{(1-\phi)(\widehat{\mu}-\mu)-\phi\widehat{\epsilon}+\widehat{\psi}_0-\psi_1}{1-\boldsymbol{\rho}(1-\phi)}
\end{aligned} \tag{25}$$

It now follows from (24) and (25) that:

$$\begin{aligned}
\overline{V}-V &= \frac{\overline{\mu}-\widehat{\mu}}{1-\boldsymbol{\rho}}+\frac{(1-\phi)(\widehat{\mu}-\mu)}{1-\boldsymbol{\rho}(1-\phi)} \\
&\quad +\frac{\overline{\psi}_0-\widehat{\psi}_0}{1-\boldsymbol{\rho}}+\frac{\widehat{\psi}_0-\psi_1-\phi\widehat{\epsilon}}{1-\boldsymbol{\rho}(1-\phi)}
\end{aligned}$$

which implies:

$$\begin{aligned}
&(1-\boldsymbol{\rho})[1-\boldsymbol{\rho}(1-\phi)]\left(\overline{V}-V\right) \\
= \;& [1-\boldsymbol{\rho}(1-\phi)](\overline{\mu}-\widehat{\mu})+(1-\boldsymbol{\rho})(1-\phi)(\widehat{\mu}-\mu) \\
&+[1-\boldsymbol{\rho}(1-\phi)]\left(\overline{\psi}_0-\widehat{\psi}_0\right)+(1-\boldsymbol{\rho})\left(\widehat{\psi}_0-\psi_1-\phi\widehat{\epsilon}\right) \\
= \;& [1-\boldsymbol{\rho}(1-\phi)]\overline{\mu}-(1-\boldsymbol{\rho})(1-\phi)\mu-\phi\widehat{\mu} \\
&+[1-\boldsymbol{\rho}(1-\phi)]\overline{\psi}_0-\boldsymbol{\rho}\phi\widehat{\psi}_0-(1-\boldsymbol{\rho})\psi_1-(1-\boldsymbol{\rho})\phi\widehat{\epsilon}
\end{aligned}$$

Rearranging (23), substituting for $V-\overline{V}$, and multiplying the resulting equation by $(\boldsymbol{\rho}-1)[1-\boldsymbol{\rho}(1-\phi)]$ yields:

$$\begin{aligned}
&(1-\boldsymbol{\rho})[1-\boldsymbol{\rho}(1-\phi)](\mu-\overline{\mu}-\epsilon) \\
= \;& \boldsymbol{\rho}[1-\boldsymbol{\rho}(1-\phi)]\overline{\mu}-\boldsymbol{\rho}(1-\boldsymbol{\rho})(1-\phi)\mu-\boldsymbol{\rho}\phi\widehat{\mu} \\
&+\boldsymbol{\rho}[1-\boldsymbol{\rho}(1-\phi)]\overline{\psi}_0-\boldsymbol{\rho}^2\phi\widehat{\psi}_0-\boldsymbol{\rho}(1-\boldsymbol{\rho})\psi_1-\boldsymbol{\rho}(1-\boldsymbol{\rho})\phi\widehat{\epsilon}
\end{aligned}$$

Collecting terms:

$$\begin{aligned}
&-[1-\boldsymbol{\rho}(1-\phi)]\overline{\mu}+(1-\boldsymbol{\rho})\mu-(1-\boldsymbol{\rho})[1-\boldsymbol{\rho}(1-\phi)]\epsilon \\
= \;& -\boldsymbol{\rho}\phi\widehat{\mu}+\boldsymbol{\rho}[1-\boldsymbol{\rho}(1-\phi)]\overline{\psi}_0-\boldsymbol{\rho}^2\phi\widehat{\psi}_0-\boldsymbol{\rho}(1-\boldsymbol{\rho})\psi_1-\boldsymbol{\rho}(1-\boldsymbol{\rho})\phi\widehat{\epsilon}
\end{aligned}$$

Given the parametric assumptions of the model we note that:

$$\overline{\psi}_0-\widehat{\psi}_0=\ln(1-\widehat{p})/(1-\overline{p})$$