Matching Job Applicants to Vacancies: An Empirical Model of Multistage Hiring

Preliminary

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

July 10, 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] See Jovanovic (1979), Miller (1984) 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: first, a subset of candidates is chosen for interviews, and then during the interview, the manager learns more about each interviewed candidate and offers one of them 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.

Our research builds on a small but growing body of literature addressing hiring processes; however, to the best of our knowledge, no prior studies have estimated a structural model utilizing 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, focusing on inter-firm competition: 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 require 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 from one firm that provides insights into their 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

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

shorter tenure and 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 manager 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 manager's. This approach introduces an agency problem, as the hiring committee must consider the manager'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 manager made all hiring decisions, thereby eliminating the divergent preferences between the committee and the manager. We find that this substantially increases the manager's value received from the hired candidate. The next set of counterfactuals explores the effects of regulatory changes on the hiring process. Specifically, we analyze how anonymizing candidates' race and gender during interview selection affects which individuals are interviewed and hired. This policy increases the share of of black candidates who are interviewed, and almost doubles the share of black candidates who are hired. In the last set of counterfactuals, we examine the role of multistage choice by comparing outcomes in alternative settings, such as random interview selection and sequential interviews, in which case only one candidate is interviewed at a time. These counterfactuals show that the candidate selection process, either done through an interview committee selecting selecting a consideration set, a manager choosing a consideration set, or a sequential interview mechanism, substantially affects who is hired and the value the manager received from the hired candidate.

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. Looking at the supply side, Babcock and Laschever (2007) document 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

[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).

work using data from a single firm include Lazear (2000), Lazear et al. (2015), and Lazear et al. (2016).

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]

When we create our estimation sample, we have to drop data on certain jobs to align with the model setup. We only keep jobs that hire one person since the current version of our model only accounts for the decision to hire 1 person.[6] Our model also requires a minimum of 2 interviewed candidates, so we drop jobs that only interview one person. For computational reasons, our model is only estimated on jobs that interview up to 11 people,

[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.

[6] The model can be modified to allow for more than 1 person to be hired for a position; we are currently working to implement this.

so we make the same restriction when creating our estimation sample.[7]

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 39,341 applicants in the data; of these, about 4.5% are African-American and a little more than half are female. About 8% of the applicants have experience working at this firm, meaning that most applicants are external to the firm. The bottom section of the table provides descriptive statistics on applications. The average person applies for almost 4 jobs over the 5 years for which we have data. Each job receives an average of almost 40 applications, but only about half are both qualified and interested in the position. On average, close to 5 candidates are interviewed for each position. We use data from 3,330 vacancies.

2.1 Search behavior

We analyze the search behavior of workers, starting with an analysis of 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.[8] 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 the prior salary of candidates who are working at the firm when they apply for a new position. After controlling for the previous year salary, we still see that women apply for lower-salary jobs than men.

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.

[7] We currently are using we use 60% of the jobs in the data. 14% of the dropped observations are due to more than 1 person being hired for the position. 12% are dropped because no one was hired for the position. 10% are dropped because only 1 person was interviewed. 5% are dropped because the position interviewed more than 11 candidates.

[8] 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.

We interpret this as the candidate learning more about the position and deciding that 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.

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 within the firm, 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 is equal to 1 if a qualified and interested candidate is interviewed for a position. Looking at the results in column (1), we see that African Americans are less likely to be interviewed. In columns (2) and (3), 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.

In Table 6, the outcome variable is equal to 1 if an interviewed candidate is offered the position. Although African-American candidates are less likely to be interviewed, there is no statistical difference in the hiring rates of black and white candidates (conditional on receiving an interview). 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.[9] 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.[10] 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 and women earn lower wages. Among other things, we 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, examine how job changes affect earnings. 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) of Table 8 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

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

[10]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 wages, and no statistically significant relationship when we only consider the qualified and interested applications. 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, 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 differentials 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 worker's decision about which job applications to pursue and the terms of employment he would accept, an interview committee's selection about whom to interview, and a manager's choice over interviewed applicants to fill a job vacancy .

The model structure is as follows. A pool of minimally qualified workers denoted by $\mathbb{A}$

are presented with the opportunity to apply for a newly vacated position. A subset of them denoted by $\mathbb{B} \subseteq \mathbb{A}$ apply for the job. Then an interview committee vets each applicant $a \in \mathbb{B}$ and selects a consideration set $\mathbb{C} \subseteq \mathbb{B}$ for interview. At the interview, the manager learns the productivity of each candidate $a \in \mathbb{C}$. Additionally, each candidate $a \in \mathbb{C}$ forms a reservation wage for accepting the new position. The manager then 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.[11] The expected current value of pecuniary and nonpecuniary 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 that a worker can infer at time t and δ is the discount rate.

Each worker stays at their current spell until they receive their next job offer. The process to begin a new position starts when they see an employment opportunity, which occurs 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 does not guarantee success since he will compete with other applicants for the position. 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 $\tau_0 = 0$ by our timing convention, and the definition of J_i implies $\sum_{j=1}^{J_i} d_{ij} \exp(-\delta \rho_{ij}) \xi_{ij} = \exp(-\delta \rho_{iJ_i}) \xi_{iJ_i}$.

In the remainder of this section we omit the i and j subscripts for notational simplicity.

[11] 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.

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 the flow benefits of his current job by $w+u$, and denote the submission cost for the new employment opportunity by ξ. The nonpecuniary amenity flow of the new job is denoted by $\widehat{u}$. If he pursues the new employment opportunity by setting $d=1$, his relocation cost ϵ 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 determine whether or not a worker submits an application for the job. The worker considers (1) $\widehat{w}-w$, the difference in wages between current employment and the advertised job, (2) $\widehat{u}-u$, the difference in their nonpecuniary benefits, (3) $e^{-\delta\rho}$, the discount factor on waiting until the next employment opportunity appears that might partially equalize these differentials, and (4) $\widehat{V}-V$, the difference in the expected lifetime utility at that future point in time. These potential benefits are weighted by ϕ, which is the probability that the worker is offered and accepts the new job. Balanced against these potential gains are $\widehat{\epsilon}$, the relocation cost of moving, which is also weighted by the probability of being hired, and ξ, the submission cost. Lemma 1, proved in Appendix B 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 the worker's reservation wage as $\overline{w}$, and denote by $\overline{V}$ denote the worker's social surplus function, or his expected lifetime utility upon taking the new job and being paid his reservation wage. 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, representing expected lifetime 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 the interview committee chooses set $\mathbb{C}$ to be interviewed by the manager for a given vacancy, and candidate $a_k \in \mathbb{C}$ is hired for this 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:

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

In this equation, M_0 is the annuity on the social surplus value to the firm from filling the vacancy, defined as the ex-ante value of the position to the firm after it becomes vacant but before the applicant pool forms, and τ_k is the random time candidate c_k quits this job spell. The manager is indifferent between $a_k \in \mathbb{C}$ and $a_{k'} \in \mathbb{C}$ if and only if the compensation offered to each of them satisfies the equality $M\left(\pi_k, w_k\right) = M\left(\pi_{k'}, w_{k'}\right)$. If their quitting times share the same probability distribution, this equality specializes to $w_k - w_{k'} = \pi_k - \pi_{k'}$, the compensating differential that would arise in a static model. 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 equation (4), if $M\left(\pi_k, w_k\right) > M_0$ then $M\left(\pi_k, w_k\right) > M\left(\pi_{k'}, w_{k'}\right)$, but if $M\left(\pi_k, w_k\right) < M_0$ then $M\left(\pi_k, w_k\right) < M\left(\pi_{k'}, w_{k'}\right)$. 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 count in a positive way.

Equilibrium hiring and wages in our model are supported by several different bargaining mechanisms that are strategically equivalent.[12] To describe one of the simplest, suppose that in the final stage of hiring the firm commits to each candidate $a_k \in \mathbb{C}$ not to offer any wage lower than the amount r_k that a_k designates at the beginning of the interview as the minimal compensation he would accept. The manager relabels the candidates so that $M\left(\pi_k, w_k\right) > M\left(\pi_{k+1}, w_{k+1}\right)$ for all $k \in \{1, \ldots, |\mathbb{C}| - 1\}$, and makes an ultimatum offer to a_1 at a wage $\widehat{w}_1$ satisfying the equality $M\left(\pi_1, \widehat{w}_1\right) = M\left(\pi_2, r_2\right)$. If a_1 rejects the offer, the manager offers a_2 a wage of $\widehat{w}_2$ where $M\left(\pi_2, \widehat{w}_2\right) = M\left(\pi_3, r_3\right)$. 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

[12]For example, the mechanism we use yields the same outcome that would occur if applicants were tested for their enthusiasm for the job in the analogue to a descending price private value procurement auction with differential scoring to account for heterogeneous quality. The personnel manager offers successively less attractive wages, keeping $M\left(\pi_c, w_a\right)$ 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, Haile and Tamer (2003).

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 $\overline{w}_k$ is therefore a dominant strategy. Because $M(\pi_k, w)$ is decreasing in w, it follows that $\widehat{w}_k > \overline{w}_k$ and hence a_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 $k \in \{1, \dots, |\mathbb{C}| - 1\}$, 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 let $\widehat{\tau}_k$ denote the optimal turnover time from the new position given the equilibrium 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, candidate c_1 accepts the manager's job offer of $\widehat{w}_1$ and

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

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, this occurs when the quality of the applicant pool is higher than expected, and the successful applicant is expected to stay with the job longer than his closest rival. The reverse occurs when $\delta M_0 > \pi_2 - \overline{w}_2$ but $E\left[e^{-\delta\widehat{\tau}_1}\right] < E\left[e^{-\delta\overline{\tau}_1}\right]$.

The Interview Committee The interview committee considers all applicants in $\mathbb{B}$ and chooses the set to be interviewed. It costs λ to interview each candidate. The committee may differentially value personal characteristics of the applicants as compared to the manager, and denote that difference in preferences for $a \in \mathbb{B}$ as $L(a)$. Conditional on $\mathbb{B}$ and any $\mathbb{C} \equiv \{a_k\}_{k=1}^{|\mathbb{C}|} \subseteq \mathbb{B}$, let $\pi_{\mathbb{C}} = \left(\pi_1, \dots, \pi_{|\mathbb{C}|}\right)$ be the set of productivities, and let $\overline{w}_{\mathbb{C}} = \left(\overline{w}_1, \dots, \overline{w}_{|\mathbb{C}|}\right)$ be the set of reservation wages. Denote by $F_{\mathbb{C}}(\pi_{\mathbb{C}}, \overline{w}_{\mathbb{C}} | \mathbb{B})$ the joint probability distribution of productivities and reservation wages for all the applicants in $\mathbb{C}$. Also let c_1 and c_2 denote the successful candidate and the runner up at the final stage, respectively. 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 $M(\pi_1, \widehat{w}_1)$, augmented by the committee's own preferences $L(c_1)$, less total interview costs $\lambda |\mathbb{C}|$. This decision process recognizes from

(5) that greater competition from the runner up c_2 lowers the winner's wages $\widehat{w}_1$ and consequently increases $M(\pi_1, \widehat{w}_1)$. 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 $F_{\mathbb{C}}(\pi_{\mathbb{C}}, \overline{w}_{\mathbb{C}})$. Formally, the committee selects $\mathbb{C} \subseteq \mathbb{B}$ to maximize:

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

In principle, the committee's optimization problem given by (6) can be solved by evaluating and comparing the $2^{|\mathbb{B}|} - 1$ potential consideration sets induced by $\mathbb{B}$, but this exercise is infeasible for large $|\mathbb{B}|$. However, the number of relevant comparison sets declines to at most $|\mathbb{B}| - 1$ if three additional conditions are satisfied. First, assume conditional on the committee's information about $\mathbb{B}$, that $(\pi_k, \overline{w}_k)$ is independently distributed for all $a_k \in \mathbb{B}$ with probability distribution function $F_k(\pi_k, \overline{w}_k \,|\mathbb{B})$. Second, assume 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, assume the committee preferences coincide with the manager's. Theorem 4 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:

$$\overline{Y}(\mathbb{C}) \equiv \int_{\mathbb{C}} \max_{a_k\in\mathbb{C}} \{M(\pi_k, \overline{w}_k)\} \, dF_{\mathbb{C}}(\pi_{\mathbb{C}}, \overline{w}_{\mathbb{C}} \,|\mathbb{B}) - \lambda |\mathbb{C}| \tag{7}$$

over $\mathbb{C} \subseteq \mathbb{B}$ if the committee's preferences are aligned with the manager's. The key difference between $Y(\mathbb{C})$ specialized to $L(a) = 0$ and $\overline{Y}(\mathbb{C})$ 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* $F_{\mathbb{B}}(\pi_{\mathbb{B}}, \overline{w}_{\mathbb{B}} \,|\mathbb{B}) = \prod_{k=1}^{|\mathbb{C}|} F_k(\pi_k, \overline{w}_k \,|\mathbb{B})$ *and suppose there exists an ordering within* $\mathbb{B}$ *denoted by* $\{a_k\}_{k=1}^{|\mathbb{B}|}$ *such that for all* $k < |\mathbb{B}|$*:*

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

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

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

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 $\overline{Y}(\mathbb{C}_N)$ 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 particular, 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 $\lambda = 0$. 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 than the manager.

Corollary 5 *Suppose there exists an ordering* $\{a_k\}_{k=1}^{|\mathbb{B}|}$ *defined by* (8), *but that* $L(c_k) \neq 0$ *for some* $k \in \{1, \ldots, |\mathbb{B}|\}$. *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* $k \in \{1, \ldots, |\mathbb{B}|\}$ *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 the interview set* $\mathbb{C}$. *There are four types of applicants labeled 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 $Y(\mathbb{C}_N)$:*

Figure 1A

$productivity\ k \backslash task\ j$	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	$\mathbb{C}^N$	$Y(\mathbb{C}^N)$
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 substitute 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 $Y(\mathbb{C}_2) = Y(\mathbb{C}_3)$ while $Y(\mathbb{C}_6) + Y(\mathbb{C}_4) > 2Y(\mathbb{C}_5)$: the gross value (before interview costs) is not, loosely speaking, concave in the size of the optimal consideration set[13], another irregularity that confounds optimization.

Example 7 (FOSD fails) *Assume the productivity of the first of two types of applicants 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 11 in Appendix B 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 $\lambda = 0$. The committee's valuations are independently drawn from uniform distributions*

[13]That is, the epigraph for the negative of the mapping formed from joining adjacent points on the graph $(N, Y(\mathbb{C}_N))$ 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. 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 jeopardize 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 consideration 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.

The preferred strategy is to solve the optimization problem as defined in equation (6), given that this requires the smallest number of assumptions. However, due to the large number of candidates in $\mathbb{B}$ for many jobs, this is not always computationally feasible since we would have to consider all possible consideration sets. In estimation, we will solve the full problem up to computational limits, and then assume the conditions in Theorem 4 hold to solve the computationally simpler problem.

4 Identification and Estimation

Our sample panel on vacancies and applications only contains five years of data. Most employees do not switch jobs in the sample frame, let alone more than once. For this reason we impose a hierarchical 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. This approximation is plausible given the institutional norms of this firm. 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 elaborates our estimation strategy.

Notation The unit of observation in the sample is a job vacancy denoted by $n \in \{1, 2, \ldots, N\}$. The set of potential applicants pool for the n^{th} vacancy is denoted by $\mathbb{A}_n$, the set of applicants is denoted by $\mathbb{B}_n$, and the consideration set is denoted by $\mathbb{C}_n$.

Parameterizing the model To facilitate identification and estimation we assume that when a new employment opportunity arises, workers do not know the identity of their rivals. Their information is limited to a description of the vacancy, their own personal socioeconomic 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 non-pecuniary preferences u_{an} for the job he currently holds by setting:

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

Similarly, his preferences for the n^{th} job opportunity are written as $\widehat{u}_{an} = \widehat{x}_{an}\gamma$. The relocation cost, $\sigma_\epsilon \epsilon_{an} + \mu_\epsilon$, is only revealed to him upon reaching the consideration set. We assume ϵ_{an} is an independent and identically distributed standard normal random variable, where μ_ϵ and σ_ϵ are mean and variance shifters that are parameters to be estimated.

At the interview stage, the manager evaluates candidates $a_n \in \mathbb{C}_n$ on the basis of (x_{an}, ζ_{an}), where x_{an} is observed characteristics and ζ_{an} is a match specific benefit the firm incurs by filling vacancy n with a_n. 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 ζ_{an} is 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}, η_{an}), where η_{an} is an independent and identically distributed standard normal random variable. We parameterize the committee's preferences as

$$L(a_n) = x_a\beta + \sigma_\eta \eta_{an}. \tag{12}$$

Three features differentiate the committee's preferences from the manager's. The committee cares about η_{an} but the manager does not. The values of ζ_{an} are learned at the interview, so they are observed by the manager but not the committee. The magnitudes of σ_η 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 $(\xi_{an}, \epsilon_{an}, \zeta_{an}, \eta_{an})$ is an unobserved vector. Estimation is in three stages. First we nonparametrically estimate incidental parameters and simulate unobserved variables. Next, we estimate the preference parameters of applicants and the productivity parameters of the firm, using an iterative process that we detail below. Then we estimate the committee's preferences from their choice of the consideration set, a computationally intensive estimator that draws upon the estimates obtained from the first stage. 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 inputs in the 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. We also estimate ρ_{an}, the arrival rate of new job opportunities,[14] and $\widehat{\tau}_{an}$, the expected duration at a new job, in the first stage. The parameter vector $(p_{an}, \rho_{an}, \widehat{\tau}_{an}, \phi_{an})$ 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.

We also simulate several components of the model. We will use the superscript s to denote each simulation draw. 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, $\eta_{an}^{(s)}$ denotes a simulated committee preference factor standing in for η_{an}.

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. 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 candidate a would pay for taking the new position, $E\left[\widehat{\epsilon}_{an}\right]$, 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. Consequently these parameters are jointly estimated.

[14]See Appendix C for details on how we estimate the arrival rates.

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 9 expresses the log odds ratio of a_n completing an application for the n^{th} vacancy.

Lemma 9 *Define:*

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

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

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

If $E\left[\widehat{\epsilon}_{an}\right]$ was known we could identify γ, the nonpecuniary benefits of amenities on the job, from equation (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, where we will use the superscript k to denote each iteration. 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. Next, form $y_{an}^{(k)}$ by substituting estimates of the ancillary parameters $(p_{an}, \phi_{an}, \rho_{an})$ 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 equation (13). An estimate of γ^{k+1} can then be found using ordinary least squares.

Productivity To estimate the productivity parameters for the firm, (α, σ_ζ), we combine data from two sources. The firm data, described in Section 2, includes information on internal candidates about their current position and wage, which we can use to estimate their reservation wage. However, this sample does not contain the wages of external candidates, hardly surprising because legal restrictions prevent the firm from using such information. Therefore we cannot use our model specification to compute the reservation wages of external candidates. We use the Survey of Consumer Expectations, which asks about individual reservation wages, to compute the distribution of reservation wages. We then use this to approximate the reservation wages for the external candidates.

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 given candidate from the consideration set 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.

Let d_{an} denote an indicator for whether $a_n \in \mathbb{C}_n$ is successful or not in the data. Write d_n as the vector of indicator variables formed from d_{an} for all $a_n \in \mathbb{C}_n$. Also let $d_{an}^{(k,s)}$ indicate whether a simulated $a \in \mathbb{C}_n$ is the successful applicant or not at the k^{th} iteration and s^{th} simulation. Recall that the winning candidate maximizes the surplus to the firm, defined in equation (4). Then we can define

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

We write $d_n^{(k,s)}(\alpha, \sigma_\zeta, M_0)$ as the vector of indicator variables formed from $d_{an}^{(k,s)}$ for all $a_n^{(s)} \in \mathbb{C}_n$. From x_n we can form z_n, which is a vector of characteristics describing the n^{th} vacancy and the composition of $\mathbb{C}_n$. At the k^{th} iteration, we minimize the distance between sample analogues to $E[d_n z_n]$ and $E\left[d_n^{(k,s)}(\alpha, \sigma_\zeta, M_0) z_n\right]$. We augment these moments with moments obtained from the wage equation.

Since the wage of the hired worker depends on the reservation wage of his closest rival for 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 simulated $\zeta_{an}^{(s)}$, the parameter values $(\alpha, \sigma_\zeta, M_0)$, along with $\left(\overline{w}_{an}^{(k,s)}, d_{an}^{(k,s)}\right)$, the simulated reservation wages of candidates in the consideration set and the induced ranking by the personnel manager, we define the closest rival to the successful candidate as:

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

Let $\overline{w}_{bn}^{(k,s)}$ denote the reservation wage of the second-best candidate $b_n^{(k,s)}$. Define:

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

From (5) and (11), when $a_n^{(k,s)}$ is offered the job , his wage is:

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

We take averages over simulation draws for each iteration k. Denote $d_n^{(k)}(\alpha, \sigma_\zeta, M_0) = E\left[d_n^{(k,s)}(\alpha, \sigma_\zeta, M_0)\right]$ and $\widehat{w}_{an}^{(k)}(\alpha, \sigma_\zeta, M_0, \sigma_\epsilon) = E\left[\widehat{w}_{an}^{(k,s)}(\alpha, \sigma_\zeta, M_0, \sigma_\epsilon)\right]$. Given reservation wages for everyone in the consideration set, at the $(k+1)^{th}$ iteration, $\left(\alpha^{(k+1)}, M_0^{(k+1)}, \sigma_\zeta^{(k+1)}, \sigma_\epsilon^{(k+1)}\right)$

minimizes:

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

where Ω is positive semidefinite conforming to $E[d_n z_n]$. Clearly $E[\widehat{w}_{an} z_n] = E\left[\widehat{w}_{an}^{(k)} z_n\right]$ and $E[d_n z_n] = E\left[d_n^{(k)} z_n\right]$. 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 $E\left[\widehat{\epsilon}_{an}^{(k,s)}\right]$ which changes the values of $\overline{w}_{an}^{(k,s)}$. Given the parameterization embodied in (10), suppose 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 $(w_{an}, x_{an}, \widehat{x}_{an})$ are data, ρ_{an} is an ancillary parameter estimated outside the model, and ϵ_{an} is simulated. In Appendix B, we 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 coefficient on nonpecuniary benefits, γ. Then (3) implies the reservation wage can be expressed as:

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

Appendix B 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 $\overline{w}_{an}^{(k,s)}$, we then compute who is offered the job using $M\left(x_{an}\alpha + \zeta_{an}^{(s)}, \overline{w}_{an}^{(k,s)}\right)$ and hence the winning $\epsilon_{an}^{(k,s)}$, which gives a simulated draw for $\widehat{\epsilon}_{an}^{(k,s)}$.

This procedure estimates the reservation wages for internal applicants, but not for external applicants because we do not observe their current wages. We use the Survey of Consumer Expectations, where workers report demographic characteristics and their reservation wages. We match applicants in our data to those in the Survey of Consumer Expectations to approximate the reservation wages of these workers.

Committee preferences The committee's preferences are estimated with a simulated methods of moments estimator (SMM). We simulate $\eta_{an}^{(s)}$, the applicant specific disturbance in the committee's utility, and draw upon estimates obtained of the applicant and the manager in the previous steps. Given a value for $(\beta, \sigma_\eta, \lambda)$, the parameters representing committee preferences, we solve for the optimal consideration sets. We construct a set of moments by interacting the characteristics of the consideration set with a series of instruments that describe the composition of the applicant pool for a job.

The most challenging part of this 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 4 hold in the applicant population, and instead search globally and compute the value of each possible consideration set. Considering only up to a maximum of 7 candidates in a consideration set, there are 900 trillion possible options. We collapse this to 9 million by excluding infeasible consideration sets using the following algorithm. Consider 2 applicants who have the same characteristics x. They differ only in η, meaning that the candidate with the lower simulated η will not be interviewed unless the candidate with the higher simulated η is also interviewed. This approach allows us to search globally for interview sets up to 7 candidates, at which point we hit computational limits.[15]

To consider interview sets with more than 7 applicants, we use an estimator that exploits a direct implication of Theorem 4. 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 similar way to the manager. In the limit they coincide. We use this marginal approach to consider consideration sets with up to 11 applicants.[16]

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 show the results of the

[15]Our approach works by generating all possible consideration sets, and then eliminating the infeasible ones. We hit computational limits when trying to generate all possible consideration sets with 8 candidates using multiple processors on a cloud computing platform.

[16]We could increase this to larger consideration sets, but given the trends in the data we do not expect it to make a substantial difference in our results.

productivity parameters, which inform the equilibrium wage determination.[17]

5.1 Model Fit

Table 10 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 4.3 interviews per position and an average of 4.9 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.

5.2 Non-pecuniary parameters

Candidates have preferences over non-pecuniary characteristics of a job. 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 11 shows the estimated values of the non-pecuniary parameters.

5.3 Committee preferences

Recall that the interview committee has preferences over candidate characteristics that may differ from the manager's preferences. We estimate the parameters β that represent the divergence of preferences over candidate characteristics. Additionally, each candidate 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 12.

5.4 Productivity parameters

Table 13 shows the parameters representing the manager's preferences. 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 M_0.

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

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 manager 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.

We first compute the optimal consideration set and chosen candidate under the baseline scenario. In the counterfactual, we set the committee divergent preferences over characteristics (β) equal to 0. The impact of this change on interviews and hiring is shown in the column labeled CF1 in Tables 14 and 15, respectively. For most characteristics, there is not a large change in who is interviewed and hired for a position. However, we see an almost doubling in the share of interviewed candidates who are black, which also translates into a large increase in the share of hired candidates who are black. Table 12 shows that the committee preferences over minority candidates was large and negative, explaining why we see this increase in minority hiring when the divergent preferences are set to 0. The bottom three rows of Table 15 show how this counterfactual impacts the value to the committee and the manager from the hired candidate. The manager valuation increases by 3% in this counterfactual. The observed committee valuation decreases by almost 20%, which is reflecting the fact that the committee preferences over characteristics are no longer considered when choosing a consideration set. In this first counterfactual, the committee preferences still play a role, as they still consider their unobserved valuation (η); recall that these terms are only observed by the committee but not the manager. These drop by less than 1% in this counterfactual, which is an unsurprising result given that the values of η are still considered when choosing a consideration set in this counterfactual.

In the next counterfactual (labeled CF2), we set the unobserved terms η to 0, while continuing to set the committee divergent preferences over characteristics β at 0. This counterfactual completely eliminates the role of the interview committee, and only the manager's preferences matter when the consideration set is chosen. This counterfactual increases the share of black candidates who are interviewed and hired, and also substantially increases the number of candidates who are interviewed. This large increase in the size of the consideration set is driven by the fact that the estimated standard deviation of the η term is relatively large, meaning that the observed values of η substantially impact the choice of consideration sets in the baseline. Since the committee cares about their preferred candidate being hired, they choose a smaller consideration set to increase the probability their favored candidate is hired (since the η terms do not impact the hiring decisions). Once the η terms are set to 0 in the counterfactual, this incentive is eliminated, and the committee chooses a larger interview set. Looking at the bottom row of Table 15, we see a 100% increase in the firm valuation in this counterfactual. This is driven by the larger consideration set: since the manager is interviewing more candidates, they see more realized productivity values, leading to a higher value from the best candidate. The observed committee preferences drop by 40%, and the unobserved committee preferences drop by over 75% since they are no longer considered when the interview set is chosen.

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 an offer, conditional on applying for the job and being minimally qualified, rather than conditional on receiving an interview?

We approach this question by implementing a race and gender blind selection of the interview set. The results of this counterfactual are shown in column CF3 of Tables 14 and 15. This policy has the greatest impact on the share of black candidates interviewed and hired. This counterfactual increases the manager's value by over 8%. Since the committee can no longer screen on race and gender, this counterfactual decreases observed committee preferences by over 15%. It also reduces the value of unobserved committee preferences by about 4%.

What is the role of gender and race duration differences? The expected duration a candidate remains on the job affects the firm's value of hiring a given candidate. In the data, women and minorities have shorter expected durations, which could potentially cause differential hiring rates. To explore the role of this factor, we equalize expected durations across race and gender to analyze how this affects hiring rates. In column CF4 of Tables 14 and 15, we set all durations to those of male and white candidates, which is the group with the largest expected duration. We see a small impact on who is interviewed and hired, but an almost 15% decrease in manager's value. Since the demographic composition of the hired candidate is similar in the baseline and this counterfactual, this result is driven by the change in the expected amount of time that a worker will remain on the job. The expected duration of female and minority candidates has increased. This result implies that the average female and minority hired candidates have lower than expected value to the manager, meaning that longer durations lower the value the manager earns. To validate this result, in column CF5 we set all durations to those of female non-white candidates, and see a large increase in firm surplus. In this case, since the average duration has decreased, the value to the firm increases due to the fact that the average hired candidate has a surplus lower than the expected value of a vacancy.

What is the role of multi-stage choice? In the last set of counterfactuals, we explore the role of the multi-stage search process to learn how it affects outcomes. We explore alternative mechanisms to understand how the selection mechanism affects who is hired, as well as the value earned 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. The results are shown in Table 16. To implement this counterfactual, we start by assuming the firm interviews 2 randomly chosen candidates. We find the winner and value of the manager's and committee's valuations from each job. We then repeat this for 3 through 7 randomly chosen candidates. The change in valuations is shown in the last column of Table 16. Starting with the case where 2 candidates are interviewed in the counterfactual, we see that this counterfactual leads to a large decrease in both manager and committee valuations. As we read down the rows, we see the reduction in manager's valuation gets smaller as we increase the number of interviewed candidates. Given that there is no interview costs in this counterfactual, this result is mechanical: as the manager interviews more candidates, he gets more draws from the productivity distribution and his expected valuation increases. Recall that in the baseline model the interview committee optimally chooses 4.3 candidates to interview. In this counterfactual, when candidates are chosen randomly, the manager's valuation is lower than in the baseline when 5 candidates are interviewed. Because the candidates are not chosen optimally, it takes more interviews to get the same surplus as in the baseline. When 6 candidates are interviewed, we see an increase in surplus as compared to the baseline.

The (sometimes very large) reduction in the manager's value is driven by two factors: (1) candidates are not chosen optimally, and (2) the number of candidates is not chosen optimally. To decompose these factors, we split the sample based on how many candidates are interviewed for a position in the baseline simulation where the interview committee chooses the consideration set. For example, consider the row in Table 16 where 3 candidates are interviewed in the counterfactual. We split the jobs into 3 groups: (1) the baseline model interviews fewer than 3 candidates, the baseline model interviews 3 candidates, and (3) the baseline model interviews more than 3 candidates. When the baseline model interviews 3 candidates, both the baseline and the counterfactual interview the same number of candidates. Therefore, for this group, the change in manager's value is only due to the non-optimal selection of interviewed candidates. In this group, we see only a 10% reduction in manager's value, a much smaller decrease as compared to the average over all jobs.

The prior counterfactual shows the outcomes under when the candidates are chosen randomly, but does not embed an optimal choice mechanism for the firm. In the last counterfactual, we combine the idea of random candidate selection within a choice framework. To do this, we assume the manager starts with an interview set with 2 randomly chosen candidates. The manager interviews both candidates, and then decides whether or not to interview a third candidate. Considering the third candidate gives them an additional draw from the distribution which could increase their surplus, but also forces the manager to pay the cost of an additional interview. The manager continues until the benefit of an additional interview is lower than the cost of an interview. See Appendix D for the derivation of the value functions used in this counterfactual. We compare the results of this counterfactual to a baseline scenario where the manager makes all hiring decisions. In this way, both the baseline and counterfactual only consider the manager's preferences, and the difference in outcomes are only due to the sequential search process. Allowing the manager to sequentially search has large impacts on outcomes. The number of candidates interviewed, on average, increases by 1.

7 Conclusion

In this paper, we leverage data on job applications from a large firm to develop and estimate a search and matching model, focusing on improving our understanding on how firm's choose whom to hire. In our model 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. Once the firm receives all applications for a vacancy, the hiring process occurs in two main stages: (1) an interview committee selects the candidates to interview, and (2) the manager chooses which of the interviewed candidates to hire. Wages are determined in equilibrium, reflecting the competition for the position.

After estimating the parameters of the firm's hiring decision process, we evaluate several counterfactual scenarios. Our counterfactuals aim to understand how hiring outcomes would vary under alternative hiring mechanisms, considering their impact on (1) the demographic composition of who is hired, and (2) the value received by the firm from the hired candidate. First, we investigate the impact on outcomes if the manager made all hiring decisions, thereby eliminating the divergent preferences between the committee and the manager. We find that this substantially increases the manager's value received from the hired candidate. The next set of counterfactuals explores the effects of regulatory changes on the hiring process. Specifically, we analyze how anonymizing candidates' race and gender during interview selection affects which individuals are interviewed and hired. This policy increases the share of of black candidates who are interviewed, and almost doubles the share of black candidates who are hired. In the last set of counterfactuals, we examine the role of multistage choice by comparing outcomes in alternative settings, such as random interview selection and sequential interviews, in which case only one candidate is interviewed at a time. These counterfactuals show that the chosen hiring mechanism substantially affects who is hired and the value the manager received from the hired candidate.

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.

Haile, P. and Tamer, E. (2003). Inference with an incomplete model of english auctions. *Journal of Political Economy*, 111(1):1–51.

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.

Miller, R. A. (1984). Job matching and occupational choice. *The Journal of Political Economy*, 92(6):1086–1120.

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

Characteristics of applicants	
Share of sample that is African American	4.46%
Share of sample that is female	55.79%
Share of sample with high school degree or less	17.95%
Share of sample with some post-secondary education	36.04%
Share of sample with a college degree	35.54%
Share of sample with a graduate degree	10.47%
Share of sample with experience at the firm	8.14%
Number of applicants	39,341
Summary statistics on applications	
Average number of applications/candidate	4.88
Average number of applications per job	37.02
Average number of applicants who are qualified and interested	20.64
Average number of people interviewed per job	4.93
Number of jobs	3,330

The sample only includes jobs which interview 2-11 people and hire 1 person.

Table 2: Number of applications submitted

	All (1)	Qualified and interested (2)
African American	1.087***	0.546***
	(0.114)	(0.0720)
Female	0.115**	0.136***
	(0.0489)	(0.0310)
Duration at firm	0.116***	0.114***
	(0.0377)	(0.0239)
Observations	52,705	52,705

Standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. The dependent variable is the number of applications submitted in a year. 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	-1029.8***	254.2	-1309.3***	-748.5
	(107.4)	(593.4)	(136.0)	(705.4)
Female	-1130.6***	-2597.6***	-1257.9***	-3012.8***
	(56.89)	(252.2)	(69.11)	(288.3)
Number of applications submitted per year	-26.05***	-191.2***	-73.61***	-219.0***
	(2.315)	(25.28)	(3.204)	(31.10)
Experience at firm	86.01*	-378.3***	173.2***	-411.2***
	(47.59)	(116.9)	(56.11)	(134.8)
Experience in division	136.0	101.0	120.5	133.7
	(104.3)	(150.1)	(114.0)	(166.1)
Worked for firm in prior year	425.5**		539.1**	
	(193.0)		(225.2)	
Worked for division in prior year	1039.9***	1697.6***	1311.7***	2052.1***
	(365.2)	(475.8)	(402.8)	(530.5)
	(53.02)	(234.7)	(63.96)	(264.7)
Previous year salary		0.211***		0.232***
		(0.00692)		(0.00794)
Fulltime position	4134.4***	2779.3**	3579.8***	1649.6
	(190.7)	(1120.9)	(214.0)	(1219.2)
Job at central location	695.5***	2071.3***	824.2***	2386.5***
	(67.25)	(302.9)	(83.88)	(346.3)
Observations	102,619	5,581	65,970	4,209

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) Not interested	(2) Not qualified
African American	0.0137***	0.0537***
	(0.00374)	(0.00515)
Female	0.00874***	0.00201
	(0.00210)	(0.00279)
Number of applications submitted per year	0.000394***	0.00293***
	(0.0000752)	(0.000104)
Number applications for job	-0.0000682***	-0.000981***
	(0.0000203)	(0.0000290)
Experience at firm	-0.000333	-0.00621***
	(0.00165)	(0.00219)
Experience in division	-0.00812**	-0.00831
	(0.00389)	(0.00572)
Worked for firm in prior year	-0.00585	-0.00992
	(0.00733)	(0.00962)
Worked for division in prior year	-0.0260*	-0.0505**
	(0.0141)	(0.0199)
Fulltime position	-0.0289***	0.0884***
	(0.00770)	(0.0107)
Job at central location	0.0497***	0.0438***
	(0.00250)	(0.00351)
Average salary of posted job	-0.00000275***	0.00000751***
	(0.000000129)	(0.000000168)
Observations	101,756	101,697

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) All	(2) Internal	(3) External
African American	-0.0504***	-0.0633**	-0.0452***
	(0.00799)	(0.0315)	(0.00772)
Female	0.0131***	0.0227	0.0109***
	(0.00374)	(0.0149)	(0.00363)
Number of applications submitted per year	-0.00488***	-0.00906***	-0.00466***
	(0.000181)	(0.000936)	(0.000176)
Number of qualified/interested candidates	-0.00395***	-0.00572***	-0.00348***
	(0.0000615)	(0.000293)	(0.0000555)
Experience at firm	-0.00146	-0.00672***	
	(0.00279)	(0.00238)	
Experience in division	0.0178***	0.0310***	
	(0.00552)	(0.00591)	
Worked for firm in prior year	0.0118	0.00828	
	(0.0119)	(0.0171)	
Worked for division in prior year	0.125***	0.161***	
	(0.0195)	(0.0263)	
Fulltime position	0.0527***	0.0917*	0.0416***
	(0.0118)	(0.0541)	(0.0113)
Job at central location	-0.0239***	-0.0374**	-0.0212***
	(0.00447)	(0.0177)	(0.00433)
Average salary of posted job	0.000660***	-0.000334	0.000749***
	(0.000201)	(0.000750)	(0.000197)
Observations	65,970	6,942	60,823

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. Salary of job is measured in 1000's of dollars. Number of candidates refers to the number of qualified and interested candidates for a given position.

Table 6: Probit regression on being hired

	(1) All	(2) Internal	(3) External
African American	-0.00385	0.0815*	-0.0136
	(0.0183)	(0.0456)	(0.0196)
Female	0.0210***	0.0290	0.0183**
	(0.00762)	(0.0195)	(0.00824)
Number of applications submitted per year	-0.0118***	-0.0181***	-0.0110***
	(0.000679)	(0.00241)	(0.000904)
Number of people interviewed	-0.0357***	-0.0382***	-0.0347***
	(0.00150)	(0.00407)	(0.00245)
Experience at firm	0.00841	0.00138	
	(0.00548)	(0.00311)	
Experience in division	-0.0108	-0.00276	
	(0.00788)	(0.00478)	
Worked for firm in prior year	-0.00499	-0.0156	
	(0.0244)	(0.0275)	
Worked for division in prior year	0.0522*	0.0392	
	(0.0313)	(0.0337)	
Job at central location	0.0272***	-0.0306	0.0397***
	(0.00905)	(0.0233)	(0.00987)
Fulltime position	-0.0140	-0.00444	-0.0150
	(0.0256)	(0.0823)	(0.0266)
Average salary of posted job	-0.000207	-0.000320	-0.000225
	(0.000367)	(0.000906)	(0.000398)
Observations	15,905	2,494	13,559

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. Salary of job is measured in 1000's of dollars.

Table 7: Durations at job

	(1) OLS	(2) 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	Application sample period		
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: Model Fit

	Share of candidates			
	Interviewed		Hired	
	Model	Data	Model	Data
Women	0.50	0.54	0.57	0.57
Black	0.022	0.030	0.033	0.029
Internal to firm	0.076	0.084	0.072	0.11
High school	0.13	0.15	0.14	0.15
Some college	0.31	0.33	0.33	0.32
College	0.39	0.38	0.39	0.39
Post college	0.17	0.14	0.13	0.14
Total number of interviews	4.28	4.93		
Wage of hired candidate			38,381.3	38,306.0

Table 11: Candidate non-pecuniary preferences

		Division			
		1	2	3	4
Constant		6,306.5	22,685.2	-2,033.5	-4,186.4
Some college	382.7	4,242.6	-5,911.7	-3,351.2	657.9
College	-1,162.4	2,359.3	-6,871.4	920.4	3,156.9
Post college	-690.9	8,909.6	-19,285.6	2,280.1	-1,972.4
Female	-2786.0	-3,909.9	-2,363.6	3,957.0	4,009.6
Black	-2,408.1	-20,879.2	-8,954.2	6,003.4	-5,184.2

The data report the division each job is in, and we estimate candidate's preferences over divisions. There are too many divisions to treat each separately, so we divided them into groups based on perceived similarity.

Table 12: Committee preferences

	Estimate
High school	-1,780.2
Some college	-1,685.9
College	-1,355.4
Female	-742.2
Black	-6641.5
Internal	3665.5
Cost of interview	416.4
Standard deviation of η	6,938.4

Table 13: Productivity parameters

	Estimate
High school	-8,037.1
Some college	-6,529.7
College	4,804.4
Female	559.3
Black	-6,517.5
Internal	9,998.8
Value of a vacancy	8,619.3
Mean of ϵ	-4,709.9
Standard deviation of ϵ	9,947.5
Standard deviation of ζ	9,953.9

Table 14: Counterfactuals: Impact on interviews

	Baseline	CF1	CF2	CF3	CF4	CF5
Women	0.50	0.51	0.58	0.53	0.52	0.51
Black	0.022	0.044	0.054	0.047	0.034	0.037
Internal to firm	0.076	0.065	0.037	0.072	0.074	0.084
High school	0.13	0.13	0.14	0.13	0.13	0.14
Some college	0.31	0.31	0.35	0.31	0.31	0.31
College	0.39	0.39	0.38	0.39	0.39	0.38
Post college	0.17	0.16	0.13	0.17	0.17	0.17
Number of interviewed candidates	4.28	4.34	7.44	4.47	4.43	3.54

Counterfactuals 1 and 2 set the divergent preferences over candidate characteristics to 0, while Counterfactual 2 additionally eliminates the role of the unobserved term η in the divergent preferences. In counterfactual 3, the interview committee cannot see the race or gender of each applicant. In counterfactuals 4 and 5, the duration terms are equalized across race and gender. In counterfactual 4, all the duration terms are equal to those of white men. In counterfactual 5, they are all the same as the estimated durations for female minorities.

Table 15: Counterfactuals: Impact on hiring

	Baseline	CF1	CF2	CF3	CF4	CF5
Women	0.51	0.51	0.58	0.54	0.51	0.51
Black	0.030	0.051	0.046	0.059	0.047	0.048
Internal to firm	0.099	0.080	0.048	0.10	0.10	0.11
High school	0.14	0.14	0.15	0.14	0.14	0.14
Some college	0.32	0.32	0.37	0.32	0.32	0.32
College	0.38	0.38	0.37	0.38	0.38	0.38
Post college	0.17	0.16	0.12	0.16	0.16	0.17
Change in manager value		3.12%	102.32%	8.41%	-14.67%	59.33%
Change in observed committee value		-19.0%	-40.0%	-15.79%	-10.15%	-6.29%
Change in unobserved committee value		-0.62%	-76.89%	-4.25%	-5.54%	13.12%

Counterfactuals 1 and 2 set the divergent preferences over candidate characteristics to 0, while Counterfactual 2 additionally eliminates the role of the unobserved term η in the divergent preferences. In counterfactual 3, the interview committee cannot see the race or gender of each applicant. In counterfactuals 4 and 5, the duration terms are equalized across race and gender. In counterfactual 4, all the duration terms are equal to those of white men. In counterfactual 5, they are all the same as the estimated durations for female minorities.

Table 16: Random interview set: Change in valuations

Number of interviews (CF)	Valuation considered	Number of interviews in baseline: Less than CF	Same as CF	More than CF	Total change
	Manager		-21.43%	-1,824.8%	-1294.5%
2	Committee (observed)		-320,519.0%	-10,372.7%	-101,576.6%
	Committee (unobserved)		-65.48%	-74.15%	-71.60%
	Manager	36.90%	-10.33%	-1,171.6%	-691.8%
3	Committee (observed)	-415,665.3%	-48,096.9%	-124.14%	-111,350.7%
	Committee (unobserved)	-85.23%	-56.04%	-78.8%	-76.9%
	Manager	57.6%	-21.4%	-848.3%	-378.5%
4	Committee (observed)	-349,827.9%	-86.1%	-141.9%	-125,280.1%
	Committee (unobserved)	-86.2%	-64.5%	-84.2%	-81.4%
	Manager	68.8%	-113.6%	-535.5%	-158.2%
5	Committee (observed)	-272,384.9%	-221.1%	-141.9%	-136,467.5%
	Committee (unobserved)	-86.5%	-78.3%	-87.4%	-85.3%
	Manager	100.6%	-42.4%	-205.8%	12.5%
6	Committee (observed)	-214,429.6%	-9,698.1%	-171.6%	-140,768.7%
	Committee (unobserved)	-88.3%	-82.3%	-89.7%	-88.0%
	Manager	206.3%	433.8%	-80.0%	142.8%
7	Committee (observed)	-205,802.9%	-150.8%	-165.5%	-152,777.6%
	Committee (unobserved)	-90.8%	-68.0%	-91.5%	-90.5%

Appendix A: Comparison with ACS/PSID data

Table 17 compares the demographics of our sample to the ACS, using only labor market participants located in the a similar geographic region.

Table 17: 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 18. 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 19.

Table 18: 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 19: 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 $G_i\left(M^{(2)}\right) = G_j\left(\overline{M}^{(2)}\right)$:

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

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.[18] 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

[18] 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.

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.

■

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:

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

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:*

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

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 $M(\pi_a, w_a)$ with $M(\pi_b, \overline{w}_b)$. Appealing to (**??**):

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

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

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

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

Substituting for τ_a and τ_b gives the result.

Lemma on probability of submission

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

which comes form

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

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:

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

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:

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

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

$(\boldsymbol{\rho}-1)\left[1-\boldsymbol{\rho}\left(1-\phi\right)\right]$ yields:

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

Collecting terms:

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

Given the parametric assumptions of the model we note that:

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

Appendix C: Derivation of mixed hazard model

We can write the hazard rate for a person with type j as

$$\lambda_j(t)=\lim_{dt\to 0}\frac{Pr_j(t<T<t+dt|T>t)}{dt}$$

Denoting $S_j(t)=Pr_j(T>t)$ and $f_j(t)$ as the pdf for t, we can then write

$$\lambda_j(t)=\frac{f_j(t)}{S_j(t)}$$

Writing $F_j(t)$ as the cdf, we know $S_j(t)=1-F_j(t)$, so $-f_j(t)=\frac{d}{dt}S_j(t)$. Then we can write

$$\lambda_j(t)=-\frac{d}{dt}\log S_j(t)$$

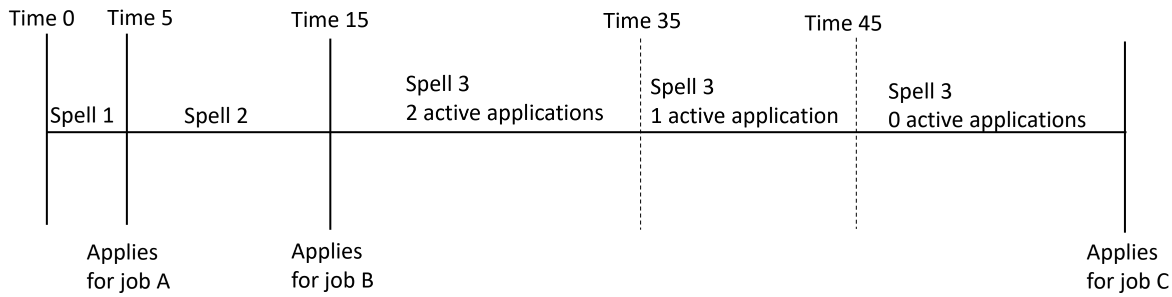
We can integrate from 0 t and impose $S_j(0)=1$ to find

$$S_j(t)=\exp\left\{-\int_0^t\lambda_j(x)dx\right\}$$

If we assume $\lambda_j(t)=\lambda_j$, then

$$S_j(t)=\exp\left(-\lambda_j t\right)$$

Figure 1: Spell example: number of current applications



The likelihood for person i, assuming they are type j, can be written as follows, where $apply_i$ indicates whether or not they applied for a position at the end of the spell

$$\lambda_j^{apply_i} S_j(t)$$

We assume that there are two types, each with a different λ_j. This allows for different types of spells: one where you are searching, and one where you maybe are not. The probability of being a type 1 is π. Then we write the likelihood for a person as

$$\pi\lambda_1^{apply_i} S_1(t) + (1-\pi)\lambda_2^{apply_i} S_2(t)$$

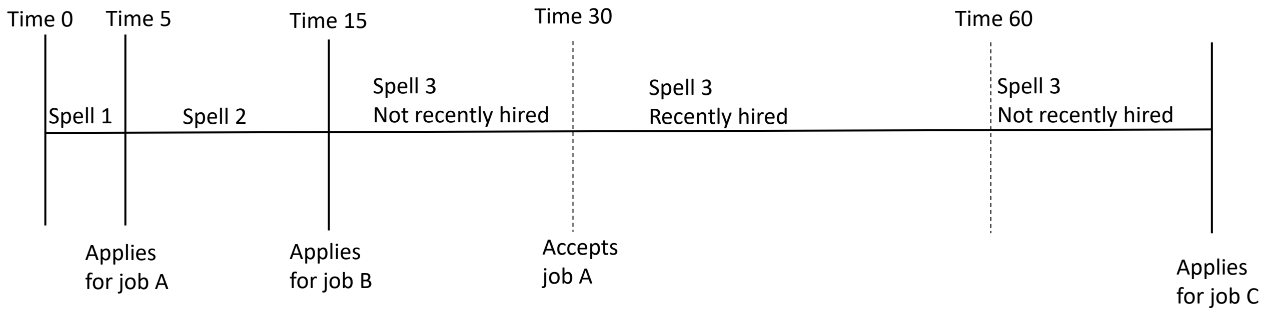
We allow λ to depend on characteristics X_t using a proportional hazards model. For $j = 1, 2$,

$$\lambda_j(X_t) = \lambda_j \exp(X_t\beta_j)$$

Note that the characteristics are time varying. Some are not, such as race, gender, and whether a person is an internal or external candidate. We also control for the number of applications within the past 30 days. This is meant to capture the fact that applications tend to happen in clusters. This can vary over the course of a spell. To see this, consider Figure 1.

In this example, a person applies for job A at time 5, job B at time 15, and job C at the terminal period. Each spell consists of the time between applications, since we are considering the hazard of submitting an application. Consider spell 3, which is the time between when a person applies for job B and C. Recall that we only consider recent applications, which we define as being submitted within the past 30 days. At the start of the spell, the person has 2 current applications (for job A and job B). However, at time 35, it now has been 30 days

Figure 2: Spell example: number of current applications



since they applied for job A, so now they only have 1 current application. By day 45 they have 0 recent applications.

We additionally control for whether a person has been hired for a job in the past 30 days. To compute this, we use the information in the data on when each job offer was accepted. We control for whether or not a person has been hired within 30 days. This again can vary within a spell, as demonstrated in Figure 2. In this figure, a candidate applies for job A at time 5, and then job B at time 15. At the time they apply for job B, they have not heard whether or not they received job A. Therefore, from time 15 to time 30 in spell 3, he has not recently been hired. At time 30, he accepts job A. From time 30 to 60, we has recently been hired. However, at day 60, it has now been longer than 30 days since he accepted the job, so he no longer has a recent job offer.

Appendix D: Derivation of sequential choice counterfactual

For each candidate i, write $v_i = \pi_i - rw_i + E\left[\exp\left(-\delta\tau\right)\right]\left(M_0 - \left(\pi_i - rw_i\right)\right)$. This is the surplus if candidate i is hired at their reservation wage. This is known for each interviewed candidate, but unknown for each candidate prior to the interview. For a position j, assume the values of v_i are drawn from the normal distribution with mean μ_j and standard deviation σ_j.

Suppose the manager has interviewed $\mathbf{I}$ candidates and learned the value v_i for each of them. Denote $\mathbb{I} = \{v_i\}_{i=1}^{\mathbf{I}}$ as the known information about each of the interviewed candidates. Denote n as the number of candidates who have yet to be interviewed. The state space for this problem is $\{\mathbb{I}, n\}$

We write $M\left(\mathbb{I}\right)$ as the value of v_i for the second best candidate in the interviewed set.

This tells us the surplus to the firm if they choose to hire from this set. Write the cost of an interview as λ.

We write the value function as follows:

$$V(\mathbb{I}, n) = \max\left\{M(\mathbb{I}), E\left[V(\mathbb{I}', n-1)\right] - \lambda\right\} \tag{26}$$