Matching Applicants to Vacancies: An Empirical Model of Multistage Hiring

Rebecca Lessem and Robert A. Miller *

June 27, 2026

Abstract

We exploit five years of data on job vacancies, applications, and hiring drawn from a large firm to understand the hiring process. We develop, identify, and estimate a dynamic noncooperative game in which hiring unfurls as a multistage process: workers choose whether to pursue new employment opportunities, an interview committee chooses which candidates to interview, and then a manager interviews each candidate and ultimately hires one. These stakeholders may have different objectives and we use the estimated parameters to evaluate the losses associated with this agency problem. Then we predict how hiring outcomes and stakeholder's benefits would change under alternative mechanisms. These include race and gender blind interview selections at the second stage, and amalgamating the last two stages into one.

KEYWORDS: multistage hiring, search and matching, agency problem, consideration set, panel data on vacancies, applications, interviews, hires and job durations within a single firm.

1 Introduction

It is remarkable how little labor economists know about how workers and firms match up. A vast literature studies workers' search behavior and the frictions that shape job finding, but the process through which firms select among applicants is less researched. There are two major roadblocks. First, data on who applies to job vacancies and how firms choose among the applicants is scarce. Second, hiring is often a multistage affair in which the consideration set of viable candidates endogenously shrinks with each stage to the point of filling the vacancy. Solving this problem is computationally demanding.

Yet understanding the hiring process is important. For example, anti-discrimination policies may influence applications, interview selection, who is offered the job to fill the vacancy, and on what terms. It is difficult to accurately predict how a policy innovation

*Tepper School of Business, Carnegie Mellon University: rlessem@andrew.cmu.edu, ramiller@cmu.edu

would affect labor market outcomes without analyzing its effects at each stage of the hiring process. Our study clears both roadblocks, enabling us to conduct counterfactuals of the sort described.

The first roadblock is essentially empirical, which we overcome with a unique data set. This paper uses data gathered from a large firm over a five year period, tracking the progress of all applicants for each vacancy. This includes an initial screening, interview selection, and job offers. We model these data as equilibrium outcomes for a dynamic game of worker search and firm hiring. The data is rich and extensive enough for us to semi-parametrically identify and parametrically estimate the primitives of our model, and thus quantify the mechanisms through which the firm selects candidates for interviews and ultimately hires workers to fill vacancies.

The second roadblock is computational. In our model, a committee selects the subset of applicants to interview from those candidates who decided to pursue the job vacancy. This is a high-dimensional combinatorial optimization problem complicated by an agency problem that arises because the preferences of the committee selecting the consideration set and the manager choosing from the selected set may not be aligned. We develop a computationally tractable solution that closely approximates optimal decisions for the moderate numbers of applicants per vacancy observed in our data set, and nest this solution within our estimator of the committee's preferences.

To motivate our formal analysis, we first document empirical patterns in search, hiring, and post-hire outcomes. African American and female applicants apply to more positions than white male applicants.[1] Black applicants are less likely to be selected for interviews, yet conditional on interviewing, they are slightly more likely to receive an offer. Post-hire outcomes also differ: both black and female employees experience shorter tenure and slower wage growth.

In our model, a pool of candidates is initially informed about a job vacancy that the firm wants to fill. Each candidate decides whether to complete an application or not. A committee reviews the applicant pool (the candidates who apply) and selects a subset of them, a consideration set, to be interviewed. The hiring manager then acquires additional information during the interview stage and chooses whom to hire. The committee's preferences may diverge from those of the hiring manager, creating an agency problem. This conflict is ameliorated by information the manager acquires about applicant quality during the interview. The wage offer to the manager's preferred candidate is determined by the intensity

[1]Empirical work typically infers search intensity from job-to-job transitions, a measure that conflates workers' search effort with their probability of receiving an offer. A notable exception is Faberman et al. (2022), which collects data on workers' search intensity and studies how it affects employment outcomes.

of competition for the position. The equilibrium of this noncooperative game provides the data-generating process underlying our econometric framework.

Our empirical analysis evaluates how alternative hiring mechanisms affect the composition of hires, the terms of employment, and the surplus generated by the firm. We find that eliminating the divergence of preferences between the committee and the manager increases both the representation of women and minority candidates among hires and the firm's expected surplus. At the same time, the committee creates substantial value by screening applicants before interviews take place. Replacing committee-based screening with random interview selection dramatically increases the interview costs required to generate the same level of surplus. Our last counterfactual is motivated by the pronounced racial and gender disparities observed in the data. When the interview committee is race- and gender-blind, the share of minority hires increases substantially, highlighting the extent to which hiring procedures shape workforce composition.

To our knowledge, no prior work has structurally estimated a model using data on firms' actual interview and hiring choices, but our research connects to several strands of the literature. The closest is the search and matching literature. Seminal contributions include Miller (1984), van den Berg and Ridder (1998), Postel-Vinay and Robin (2002), Cahuc et al. (2006), Flinn (2006), Bagger et al. (2014), Lise et al. (2016), Bagger and Lentz (2019), Lentz et al. (2023), and Lentz and Roys (2024). We build on this work by introducing a framework that models how firms select whom to interview and ultimately hire from a given pool of applicants; this is a stage of the matching process that is central in practice but largely abstracted away in existing models.

Some recent work analyzes interview and hiring decisions within unified theoretical frameworks, but none structurally estimate these decisions using data on firms' actual interview choices. Gottardi et al. (2025) develop and calibrate a model in which firms use observable signals to allocate scarce interview slots, showing how screening can magnify group differences in equilibrium. Their counterfactuals examine how interview composition and hiring rates shift when particular signals are removed. Vohra and Yoder (2024) propose a related theoretical model of interview choice that emphasizes inter-firm competition: a firm may decline to interview a candidate it expects will receive an offer from a more desirable employer.

Our treatment of selecting candidates for interview connects directly to the growing literature on consideration sets, which examines how agents construct the subset of options they evaluate when faced with a large choice environment. Barseghyan et al. (2021) develop a methodology for discrete choice settings in which the true choice set is unobserved (unlike our framework), and Coughlin (2025) applies this framework to health insurance plan selection.

Similarly, Goeree (2008) develops and estimates a model where consumers do not see the entire choice set (a contrast to our model), which can be expanded through advertising.

A large body of research documents persistent wage gaps by race and gender.[2] Here, we focus on how the search and matching process contributes to these disparities. Several empirical studies investigate links between demographic characteristics and interview and hiring decisions.[3] Our framework quantifies how changes in hiring regulations, such as anonymizing demographic characteristics during interview selection, affect interview and hiring rates.

A limitation of our work is that our estimates speak directly to the applicants, employees and management of one firm. Nevertheless, this particular setting is well suited for studying the search and matching process: the firm has a large stock of long-tenured employees, and applicants for new positions include both internal and external candidates, providing rich variation in search behavior and hiring decisions. Economists have long relied on single firm datasets to answer questions of empirical importance. Hoffman et al. (2018) study how the quality of the hired candidate is affected by whether or not a manager follows test-score recommendations. Ichnioski et al. (1997), Lazear (2000), and Bandiera et al. (2007) examine how human resource practices shape productivity. Hamilton et al. (2003), Mas and Moretti (2009), Lazear et al. (2015), and Hoffman and Tadelis (2021) use firm level data to study the role of teams and managers. Like these studies, our research is fundamentally tied to detailed, proprietary information that only individual firms can provide.[4]

2 Data

Our data come from a large, anonymous firm employing roughly 65,000 workers. Over a five-year period in the first decade of the 21st century, we observe every vacancy posted by the firm and every application submitted. Anticipating the computational costs incurred when solving the committee's problem of selecting a consideration set, we retain only those vacancies where between two and eleven applicants are interviewed, and either one or two of them are hired; that is 68% of all positions.[5]

We observe basic demographic characteristics of job candidates, including race, gender,

[2] For racial wage differentials, classic and recent evidence includes Altonji and Blank (1999), O'Neill (1970), O'Neill (1990), Neal and Johnson (1996), Carneiro et al. (2005), and Golan et al. (2024). The gender wage gap is surveyed in Blau and Kahn (2017), with additional contributions from Gayle and Golan (2012), Gayle et al. (2012), and Xiao (2021).

[3] See, for example, Bertrand and Mullainathan (2004), Shukla (2024) and Russo and van Ommeren (1998).

[4] There are many more examples of research using data from only one firm, including Baker et al. (1994a), Baker et al. (1994b), Bandiera et al. (2010), Kahn and Lange (2014), and Lazear et al. (2016).

[5] We drop 8% because no one was hired, 11% because only one applicant was interviewed, 10% because more than eleven of them were interviewed, and 3% because more than two were hired.

and age (by decade), as well as their progress through the hiring process. It unfolds in several stages. Some candidates disengage, classified here as *not interested*; this includes outcomes such as "Applicant not interested", "Failed to respond to HR", and similar labels. Others are screened out because they lack the required qualifications, and the firm designates these candidates as *not qualified*. Among the remaining (qualified and interested) candidates, which we call *applicants*, the data record who is interviewed, who receives an offer, compensation, and whether they accept it. The data set also contains information about each job, including its administrative division, and a short description of the position. Identifiers link applicants to specific vacancies and track their behavior across all applications submitted during the five-year window.

We supplement our primary data set with additional observations on annual wages and their division for every employee in the organization over a period of more than two decades. Merging the application records with these wage histories, we observe post-hire wage trajectories, as well as job and firm tenure.

Table 1 reports summary statistics on candidates. Over the five-year period, the dataset includes 39,482 individuals. Approximately 4.5% are African American, and a little more than half are female. Slightly more than half the candidates are less than 40 years old. Roughly 8% have prior experience working at the firm; the vast majority apply from outside the organization. To gauge how well the firm represents the broader labor market in the region, the second column in Table 1 shows comparable statistics from the American Community Survey (ACS) of people in a similar geographic region in similar years. Our sample has more women and fewer minorities than the ACS, as well as higher education levels. It is also younger, not surprising since the incentives for job search are larger at younger ages. The column labeled *ACS** further refines the ACS sample to include only those workers employed in industries similar to those in our firm.[6] Controlling for selection on industry shrinks the distance between the ACS and our primary sample in the distribution of education, suggesting that our data is informative about a large, albeit not exhaustive, sector of the workforce.

In our model, the reservation wage for taking a new job depends on one's current wage; however, our primary sample only includes the current wages for internal applicants but not for external ones. For this reason we also draw on the Survey of Consumer Expectations (SCE), which asks workers about their reservation wages (Federal Reserve Bank of New York, 2024). The two right columns of Table 1 show the distribution of education and age in

[6] These industries include: Information, Finance and Insurance, Professional, Scientific and Technical Services, Management of Companies and Enterprises, Administrative and Support and Waste Management and Remediation Services, Educational Services, and Public Administration.

Table 1: Characteristics of candidates

	Primary sample	ACS	ACS*	SCE	SCE*
African American	0.045	0.063	0.065		
Female	0.56	0.48	0.56		
High school degree or less	0.18	0.47	0.29	0.11	0.079
Some post-secondary education	0.37	0.25	0.23	0.34	0.29
College or graduate degree	0.45	0.27	0.48	0.54	0.63
College degree	0.35	0.18	0.28		
Graduate degree	0.10	0.09	0.20		
Average age	37.97	42.31	43.23		
Age under 40	0.56	0.42	0.40	0.28	0.41
Age 40-60	0.42	0.47	0.49	0.40	0.44
Age over 60	0.025	0.11	0.11	0.32	0.15
Experience at the firm	0.08				
Number of observations	39,482	1,727,612	484,639	9,965	1,498

The unit of analysis is an applicant in our data or a respondent in the ACS or SCE. For race, gender, education, and age we show share of candidates with given characteristics. We record the share of the sample with experience at the firm. The ACS statistics only include individuals living in the same region as our firm. The ACS* subsample limits to people in industries most similar to the firm used in this paper. The SCE* subsample only includes respondents who are searching for a job.

the SCE sample. The column labeled *SCE* uses all observations in a similar industry as our data, and the column labeled *SCE** restricts to respondents who are actively job-searching. Compared to our data, the SCE samples from a group with a higher level of education.

Turning to current wages, Table 2 compares the average annual earnings across the samples in 2010 US dollars. The first column shows the average wage of all workers in the firm. They earn substantially more than the subset of job seekers, displayed in the second column, but their wages are quite close to those of the restricted ACS sample, which is drawn from comparable industries, and also the SCE average. Another noteworthy feature is that job searchers in both the firm and the SCE earn less than their respective counterparts who are not currently engaged in job search. In summary, while the firm we investigate is not truly representative of the overall workforce, the mass of the characteristics and wages of its employees are similar to the workforce population.

Table 2: Average annual earnings

Firm	Firm*	ACS	ACS*	SCE	SCE*
44,542	34,826	35,536	42,496	41,631	40,510
(29,941)	(19,281)	(42,243)	(46,692)	(70,663)	(51,618)

The cell entries in the firm column refer to the firm's employees. Firm* is only job-searchers within the firm. ACS* is individuals in the ACS working in industries similar to the firm used in this paper. SCE* is the earnings of respondents in the SCE who are actively job searching. Standard errors in parentheses.

Search behavior The hiring process is wedged between search initiated by the firm and job seekers, and the length of the job spell. In addition to productivity and wages, the firm compares the expected duration of a new hire when comparing competing applicants; when an employee quits, the firm restarts a renewal problem. Along with differentials in wages and nonpecuniary benefits, prospective hires consider future employment opportunities. To place our analysis in context, we trace through some empirical patterns that describe multistage hiring at the firm, briefly describing aspects of the worker's search that precedes it and the job duration times that end it. All three play an important role in shaping behavior.

Candidates initiate job applications in clusters, often submitting several applications within a short period, pausing if unsuccessful, only to resume applying later. Within 10 days of submitting a job application, 31% of the candidates apply for another job. Even candidates who do not apply again within the 10 day window often apply in the future. Among applicants who do not submit another application within 10 days, 69% eventually apply again.[7] To capture this submission process, we estimate a mixed hazard model. The model assumes two levels of latent search, intensive and cursory. Search behavior is fixed within a spell but redrawn at the start of each new spell. The estimates generate the patterns observed in the data: intensively searching individuals submit multiple applications in rapid succession, while cursory searchers exhibit a much lower hazard rate. Over time, individuals cycle between these latent states, producing clustered bursts of application activity followed by periods of inactivity.

Table 3 reports the parameter estimates from the mixed hazard model. The probability of iniating job applications varies with gender, race, internal versus external status, the number of recent applications, and whether the individual was recently hired. The estimated hazard of cursory search is effectively zero for all observed characteristics. Only intensive search yields new employment opportunities. Females and minorities exhibit higher hazard rates for initiating applications.

Review process After initiating an application, roughly 14% of candidates do not complete it, and these cases are coded as *not interested* in the data. We interpret this outcome as the candidate learning more about the position during the application process and concluding that the job is not a good fit.

The review process unfolds in several stages. Candidates are screened for minimal qualifications, a step we refer to as being *qualified*. The applicant pool comprises interested candidates who meet the screening criteria. From this pool, a subset of applicants is selected for interview, and ultimately one of them receives a job offer. The selected person then

[7] Adjusting the 10 day window to 30 days yields similar results.

Table 3: Hazard model parameter estimates

Level of search	Cursory	Intensive
Baseline hazard	0.00035	0.017
	(0.0000500)	(0.000092)
Female	-0.0077	0.064
	(0.014)	(0.0062)
Black	0.33	0.23
	(0.026)	(0.012)
Internal	-0.41	-0.074
	(0.027)	(0.010)
Number applications	1.05	0.27
	(0.0022)	(0.00085)
Recently hired	0.018	-0.11
	(0.17)	(0.031)
	Probability of cursory search	0.40
		(0.0014)

Standard errors in parentheses.

chooses to accept or decline the offer.

Table 4 shows summary statistics at the application and vacancy level from our primary data set. On average, candidates initiate nearly five applications over the five year period, although the median number is only two. Each vacancy attracts close to 40 candidates; about three-quarters of them are qualified and interested. On average, five applicants are interviewed for each job.

Table 4: Applications and interviews

	Average	Median	Standard deviation
Number of applications per candidate	4.73	2	9.86
Initiated applications per job	38.96	26	42.53
Qualified and interested applications per job	28.86	12	28.86
Interviews per job	5.01	5	2.32
Number of jobs	3,684		

To examine the determinants of this culling, we estimate four probit regressions to analyze each stage of the application process. Each regression includes controls for race, gender, tenure in the firm and division, and the posted salary, as well as year, division, and occupation fixed effects. Given our focus in the counterfactual exercises, here we report by row in Table 5 the marginal effect of race and gender at different stages of the hiring process.[8]

The dependent variable relating to the first row is an indicator variable for whether the candidate becomes uninterested. The estimates show that African American candidates and

[8]The estimated regression coefficients and their asymptotic standard errors are available upon request.

women are more likely to begin applications for positions they ultimately decide not to pursue. The second row of Table 5 comes from a probit regression in which the outcome equals 1 if a candidate is not qualified for the position, which occurs in 24% of applications. The estimates show that African American candidates are more likely to apply for positions for which they do not meet the minimum qualifications.

The third row presents marginal effects from a probit regression where the outcome equals 1 if a qualified and interested candidate is selected for an interview. Here, we find that African American candidates are less likely to be interviewed. The bottom row of Table 5 show the marginal effects from a probit regression where the outcome equals 1 if an interviewed candidate receives a job offer. Although African American candidates are less likely to reach the interview stage, conditional on being interviewed there is no statistically significant difference in offer rates between black and white candidates.[9]

Table 5: Race and gender impacts on the application process

Dependent variable	Black	Female
Not interested	0.0129	0.010
	(0.0035)	(0.0020)
Not qualified	0.0461	0.0008
	(0.0043)	(0.0024)
Interviewed	-0.0454	0.0106
	(0.0069)	(0.0033)
Hired	0.0102	0.0193
	(0.0174)	(0.0075)

Each row reports the marginal effects on gender and race from a probit regression with different outcome variables. We control for education, age, search behavior, number of applications for the position, work experience, and the salary of the position, and include year, division, and occupation fixed effects. Standard errors in parentheses.

Job outcomes We wrap up the description of our data by summarizing how wages and turnover are affected by demographic characteristics and experience. Table 6 reports the results of the regressions on the logarithmic annual earnings. Column (1), computed from the primary and supplementary data for all the firm's employees, shows that African American workers and women earn lower wages, even after controlling for experience and other covariates. We include an indicator for switching divisions in any given year because it is a good proxy for job turnover within the firm. The coefficient is positive, consistent with the hypothesis that internal job turnover leads to higher earnings.

[9] We also examined how likely a job offer would be accepted. However, this exercise did not reveal any statistically significant patterns, largely because more than 95% of offers are accepted. Given the limited variation and lack of informative results, this probit is not reported.

To examine more closely how job changes affect wages, columns (2) through (4) are computed off the primary sample alone in order to directly exploit data on the timing of job transitions. The results in column (2) corroborate the findings of column (1): internal job turnover yields higher wages. Columns (3) and (4) are also based on the primary sample. Controlling for the number of applications submitted in the prior year, we test whether job search predicts wages. Column (3) reports a small negative and weakly statistically significant relationship between the number of applications submitted and earnings. When the number of qualified and interested applications is included as the regressor measuring job search, an insignificant coefficient results, as can be seen from column (4). While Postel-Vinay and Robin (2002) argue that on-the-job search raises wages through improved bargaining power with the current employer, we find no evidence of a positive relationship between wages and internal job search.

Table 6: Log earnings regression

	(1)	(2)	(3)	(4)
	Firm	Application in 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***
	(0.0061)	(0.0125)	(0.0228)	(0.0228)
Duration in current division	0.0960	0.209	0.118	0.118
	(0.0017)	(0.0040)	(0.0089)	(0.0089)
Experience squared	-0.00339	-0.00900	-0.00421	-0.00420
	(0.0001)	(0.0002)	(0.0005)	(0.0005)
Switches divisions	0.144		0.0818	0.0796
	(0.0161)		(0.0631)	(0.0632)
New job		0.216		
		(0.0172)		
Number applications submitted			-0.0048	
			(0.0026)	
Number qualified/interested submitted				-0.0055
				(0.0038)
Constant	9.772	9.176	9.511	9.510
	(0.0279)	(0.0409)	(0.0755)	(0.0756)
Observations	50,524	15,790	3,153	3,153

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. Standard errors in parentheses.

Aside from wages and productivity, job duration is potentially an important component of the firm's selection process because managers may factor expected tenure into their hiring decisions. Table 7 analyzes the relationship between job duration and a candidate's race, gender, and other personal characteristics, as well as features of the job itself.[10] The first column presents OLS estimates, while the second reports results from a Tobit model that account for right-censoring of spell durations in the last observed job spell. The results show that both African American workers and women have shorter job durations than white males. This pattern is consistent with findings from Gayle et al. (2012), who document similar results in a very different context, executives working for large firms.

Table 7: Job duration

	OLS	Tobit
African American	-0.929	-1.189
	(0.215)	(0.288)
Female	-0.295	-0.754
	(0.101)	(0.135)
Year job started		0.810
		(0.00994)
Observations	11,347	11,347

Let $y = X\beta + \gamma t_0 + \varepsilon$, where y is job duration, X is characteristics, t_0 is the start year and T_f is the final year of available data. The dependent variable in the OLS regression is $y^* = \min\{y, T_f - t_0\}$. In the Tobit specification we assume ε is normally distributed but account for the fact that y^* rather than y is observed. Both regressions control for education, unemployment and labor market participation in a given year, along with division and occupation. Standard errors in parentheses.

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 candidates denoted by $\mathbb{A}$ are presented with the opportunity to apply for a newly vacated position and initiate the application. A subset of them, the applicants, 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 his valuation of each applicant $a \in \mathbb{C}$. Additionally, each applicant $a \in \mathbb{C}$ forms a reservation wage for accepting the new position. The manager then makes as many job offers as is necessary to fill the position.

[10] The number of consecutive years a person works within a given division measures duration.

Workers New jobs start at discrete intervals in time. 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. 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 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 and initiate an application, 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 complete the application 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:

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

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. Suppose the worker is currently facing a new employment opportunity, and let $d \in \{0, 1\}$ denote the worker's decision to complete the 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 $\widehat{u}$. 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 completes 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

[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.

that future point in time. These potential benefits are weighted by ϕ, the probability that the worker is offered and accepts the new job. Balanced against these potential gains is ξ, the submission cost. Lemma 1 shows how the components are linked.[12]

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

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

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:*

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

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.

Manager Suppose the interview committee chooses set $\mathbb{C}$ to be interviewed by the manager for a given vacancy, and applicant $a \in \mathbb{C}$ is hired for this position at time $t=0$. Let π_a denote the manager's valuation of this candidate, which we will refer to as the applicant's *productivity*, which nets out relocation and learning costs. Suppose he is paid wage w_a determined upon hiring. The annuity value to the manager, which we will refer to as the *manager surplus*, at the time he is hired is:

$$M\left(\pi_a, w_a\right)=\pi_a-w_a+E\left[e^{-\delta\tau_a}\right]\left[M_0-\left(\pi_a-w_a\right)\right]. \tag{4}$$

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_a is the random time candidate a quits this job

[12] Proofs for all theorems and lemmas in the paper are in Appendix A.

spell. The manager is indifferent between $a \in \mathbb{C}$ and $a' \in \mathbb{C}$ if and only if the compensation offered to each of them satisfies the equality $M(\pi_a, w_a) = M(\pi_{a'}, w_{a'})$. If their quitting times share the same probability distribution, this equality specializes to $w_a - w_{a'} = \pi_a - \pi_{a'}$, the compensating differential that would arise in a static model. Alternatively, set $w_a - w_{a'} = \pi_a - \pi_{a'}$, but suppose $E\left[e^{-\delta\tau_a}\right] > E\left[e^{-\delta\tau_{a'}}\right]$. From (4), if $M_0 > \pi_a - w_a$ then $M(\pi_a, w_a) > M(\pi_{a'}, w_{a'})$, but if $M_0 < \pi_a - w_a$ then $M(\pi_a, w_a) < M(\pi_{a'}, w_{a'})$. 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.[13] To describe one of the simplest, suppose that in the final stage of hiring the firm commits to each candidate $a \in \mathbb{C}$ not to offer any wage lower than the amount r that a designates at the beginning of the interview as the minimal compensation he would accept. The manager relabels the candidates so that $M(\pi_a, r_a) > M(\pi_{a+1}, r_{a+1})$ for all $a \in \{1, \ldots, |\mathbb{C}| - 1\}$, and makes an ultimatum offer to a_1 at a wage $\widehat{w}_1$ satisfying the equality $M(\pi_1, \widehat{w}_1) = M(\pi_2, r_2)$. If a_1 rejects the offer, the manager offers a_2 a wage of $\widehat{w}_2$ where $M(\pi_2, \widehat{w}_2) = M(\pi_3, r_3)$. This continues until a candidate accepts her offer, or everybody rejects their respective offers, in which case the firms pays a penalty to restart the hiring process at some future date by drawing a new set of applicants. The information the applicant reports to the manager helps determine whether he is offered the job, but not his wage conditional on being offered the job. Revealing his reservation wage $\overline{w}_a$ is therefore a dominant strategy. Because $M(\pi_a, w)$ is decreasing in w, it follows that $\widehat{w}_a > \overline{w}_a$ and hence a would accept the manager's offer.

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

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

[13] 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 $M(\pi_a, w_a)$ 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).

On the equilibrium path, candidate a_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 $\pi_2 - \overline{w}_2 > M_0$ 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 $M_0 > \pi_2 - \overline{w}_2$ but $E\left[e^{-\delta\widehat{\tau}_1}\right] < E\left[e^{-\delta\overline{\tau}_1}\right]$.

Committee The committee selects a subset $\mathbb{C} \subseteq \mathbb{B}$ to interview from the applicant pool $\mathbb{B}$. It costs λ to interview each candidate if more than one is interviewed. The committee may differentially value personal characteristics of the applicants as compared to the manager, and we denote these divergent preferences as $L(a)$. Also let a_1 denote the successful candidate, and a_2 the runner up at the final stage. The committee maximizes the expected surplus $M(\pi_1, \widehat{w}_1) = M(\pi_2, \overline{w}_2)$, augmented by the committee's own preferences $L(a_1)$ for the winner less total interview costs $\lambda(|\mathbb{C}| - 1)$. The committee does not observe productivity or reservation wages, since these are revealed at the interview stage, so it integrates over the unobserved characteristics of all possible combinations of the two top candidates in any given $\mathbb{C}$. Formally, the committee selects $\mathbb{C} \subseteq \mathbb{B}$ to maximize:

$$\mathbf{L}(\mathbb{C}) \equiv E_{\mathbb{B}}\left\{\sum_{(a',a'')\subseteq\mathbb{C}} \mathbf{1}\{(a', a'') = (a_1, a_2)\}\left[L(a') + M(\pi_{a''}, \overline{w}_{a''})\right]\right\} - \lambda(|\mathbb{C}| - 1),$$

where the expectation operator $E_{\mathbb{B}}[\cdot]$ conditions on all the information available to the committee about the application pool $\mathbb{B}$.

Computing the equilibrium consideration set In principle, the committee's optimization problem 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 $|\mathbb{B}|$. A simple rule for determining the consideration set is to add candidates, one at a time, maximizing the value of the consideration set with each additional candidate, until adding another does not increase the set's value net of the cost. This reduces the number of comparison sets to at most $(|\mathbb{B}| - 1)(|\mathbb{B}| + 2)/2$. An even simpler rule is to rank all the candidates before selecting the consideration set, and then follow the ranking by adding candidates to the consideration set up to the point where the marginal increment in $\mathbf{L}(\mathbb{C})$ is less than the interview cost. This reduces the number of comparison sets to at most $|\mathbb{B}|-1$.

When are these simple rules optimal? We assume the information the manager acquires is independently distributed across candidates, meaning $(\pi_a, \overline{w}_a)$ is independently distributed for all $a \in \mathbb{B}$. To answer the question, we focus first on situations where the preferences of the committee coincide with those of the manager. This might happen when the manager exercises lots of discretion over the composition of the committee. We provide sufficient conditions for optimally including an applicant from the consideration set, and applying these conditions to all the applicants generates the simplest rule as the solution to choosing the consideration set. Then we investigate, here briefly, in more detail in our application, the usefulness of both rules when the committee's preferences diverge from their manager.

When two applicants can be ranked by first order stochastic dominance (FOSD), as formally defined in the next lemma, then both are included (excluded) in the consideration set if lower (higher) ranked applicant is.

Lemma 3 (FOSD) *Denote by $F_a(\pi_a, \overline{w}_a\,|\mathbb{B})$ the probability distribution function of $(\pi_a, \overline{w}_a)$ conditional on the committee's information about the applicant pool $\mathbb{B}$. Suppose that for $a' \in \mathbb{B}$ and $a'' \in \mathbb{B}$:*

$$\int\int_{M(\pi_{a'}, \overline{w}_{a'}) \leq M} dF_{a'}(\pi_{a'}, \overline{w}_{a'}\,|\mathbb{B}) \leq \int\int_{M(\pi_{a''}, \overline{w}_{a''}) \leq M} dF_{a''}(\pi_{a''}, \overline{w}_{c''}\,|\mathbb{B}) \tag{6}$$

for all M. Suppose $L(a) = 0$ and denote by $\mathbb{C}^o$, the optimal consideration set. If $a'' \in \mathbb{C}^o$ then $a' \in \mathbb{C}^o$.

If all the applicants can be ranked by FOSD, the simplest rule is optimal because the solution to the manager's bargaining problem also solves a concave social surplus optimization problem in which $\mathbb{C} \subseteq \mathbb{B}$ is selected to maximize:

$$\overline{\mathbf{L}}(\mathbb{C}) \equiv E_{\mathbb{B}}\left[\max_{a\in\mathbb{C}}\left\{M(\pi_a, \overline{w}_a)\right\} - \lambda\,|\mathbb{C}|\right]. \tag{7}$$

The difference between $L(\mathbb{C})$ specialized to $L(a) = 0$ and $\overline{L}(\mathbb{C})$ is that the former is the manager's criterion function, whereas the latter is a social surplus function awarding all the surplus to the firm.

Theorem 2 *Suppose $L(a) = 0$ for all $a \in \mathbb{B}$, and assume* (6) *holds for all pairs $a' \in \mathbb{B}$ and $a'' \in \mathbb{B}$. Then there exists a (reverse) preference ordering of $a \in \mathbb{B}$ denoted by $\preceq$ such that $a \preceq a+1$ for all $a < |\mathbb{B}|$. Let $\mathbb{C}_N = \{a\}_{a=1}^{N}$. If $\overline{\mathbf{L}}(\mathbb{B}) \geq \overline{\mathbf{L}}\left(\mathbb{C}_{|\mathbb{B}|-1}\right) + \lambda$ then $\mathbb{C}^o = \mathbb{B}$.*

Otherwise $\mathbb{C}^o = \mathbb{C}_{N^o}$ *where* N^o *is the unique solution to:*

$$\overline{\mathbf{L}}\left(\mathbb{C}_{N^o}\right) \geq \max\left\{\overline{\mathbf{L}}\left(\mathbb{C}_{N^o-1}\right) + \lambda, \overline{\mathbf{L}}\left(\mathbb{C}_{N^o+1}\right) - \lambda\right\}. \tag{8}$$

To illustrate why Theorem 2 fails if (6) is violated, or the preferences of the manager and the committee about applicants diverge, we consider a 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 ($\epsilon = 0$).

Lemma 4 (FOSD fails.) *There are two types of applicants, denoted by* $k \in \{1, 2\}$*. The marginal value of the first is independently and uniformly distributed on* $[\underline{\pi}, \overline{\pi}]$ *where* $\underline{\pi}+\overline{\pi} > 1$ *and* $\overline{\pi} < 1$*. The second is independently and uniformly distributed on* $[0, 1]$*. Denote the composition of the consideration set by* $\mathbb{C} = \{N_1, N_2\}$ *where* N_k *is the number of applicants selected from type* k*. Then* $\mathbb{C}_N = \{N, 0\}$ *for* $N = 2$ *and* $\mathbb{C}_N = \{0, N\}$ *for all* N *above a finite threshold that depends on* $\overline{\pi}$*.*

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, $\mathbb{C}_N$ is not monotone increasing in N. Candidates included in the optimal consideration when λ is high are excluded when λ is lower.

There are two reasons a committee includes more than one applicant, their favorite, in the consideration set. First, if the value of information revealed at the interview to the committee is affiliated with its value to the manager, then delegating some of selection process to the manager could benefit the committee. The second reason is strategic, intrinsically unrelated to information acquisition. If only one worker is included in the consideration set, then he will extract all the surplus associated with the job match, costly to the committee if its preferences are somewhat aligned with the manager's. Both factors overturn the guidance FOSD offers (when it exists) in forming consideration sets. In the following example both parties rank applicants by FOSD exactly the same way, but have different preferences.

Lemma 5 (Divergent preferences) *There are three types of applicants, denoted by* $k \in \{1, 2, 3\}$*. At the time of selecting the consideration set the manager values the* k^{th} *type of applicant as an independent uniform random variable with support* $[0, \overline{\pi}_k]$*, where* $0 < \overline{\pi}_1 < \overline{\pi}_2 < \overline{\pi}_3 = 1$ *and* $\overline{\pi}_3 - \overline{\pi}_2 = \kappa$*. Both parties value the first and third types the same way. However the committee values the second type as an independent uniform random variable* *with support* $[0, \overline{\pi}_1 + \kappa]$*. There are* $N-1$ *applicants of the third type, and one each of the other two types. If it is optimal to interview* N *applicants, the manager's preferred consideration set includes all* $N-1$ *type three applicants and the type two applicant. For sufficiently small* κ *the committee also selects this set if* $N = 3$ *, but if* $N > (6 + \pi) / (2 - \pi)$ *the committee selects all the type three applicants and the type one applicant.*

When N is small, strategic concerns are paramount. Even though the committee in this example ranks the mid-level applicant barely above the worst and knows the manager would rank him barely below the best, it includes the mid-level candidate in the consideration set. The committee hopes the mid-level candidate is the second choice of the manager, and is effective in reducing the payout to the winning applicant. It optimally runs the risk that the mid-level candidate is offered the job, much to its chagrin, beating good but unexceptional applicants whom the committee ranks more highly both exante and expost. When there are many top ranked candidates to include in the consideration set these strategic concerns are mitigated. Their rivalry shrinks the portion of the job match surplus paid to the winner. So the committee optimally eliminates their risk exposure by selecting a worse ranked candidate they know the manager is unlikely to pick, in the process violating the FOSD rule the manager would prefer. Thus the example illustrates another quirk of multistage hiring: both parties ranking applicants by FOSD the same way does not ensure the committee will act in the manager's interests when selecting the consideration set.

4 Identification

The outcomes of the data generating process comprise the pool of candidates $\mathbb{A}$, including (w_a, d_a), the current wage and application decision of each candidate $a \in \mathbb{A}$, the applicant pool $\mathbb{B} \subseteq \mathbb{A}$, a vector of observed characteristics describing each candidate x_a, the applicants selected for interview, $\mathbb{C} \subseteq \mathbb{B}$, the identity and wage of the new hire $(a_1, \widehat{w}_1)$, and his tenure upon resigning from his current job τ_1. The sample consists of N observations of vacancies given by:

$$\left(\mathbb{A}_n, \{w_{an}, x_{an}, d_{an}\}_{a_n \in \mathbb{A}_n}, \mathbb{C}_n, a_{1n}, \widehat{w}_{1n}, \tau_{1n}\right),$$

where $n \in \{1, \ldots, N\}$ reflects the dependence of each observation on a description of the worker's current job. Thus $\mathbb{A}_n$ defines details about the advertised job for the n^{th} observation along with the number of candidates who consider completing the application for the job, (w_{an}, x_{an}, d_{an}) are their characteristics including their current wage and whether they complete the application, $\mathbb{C}_n$ defines the committee's consideration set for that observation, and $(a_{1n}, \widehat{w}_{1n}, \tau_{1n})$ gives the identity of the winning candidate, his or her wage at the new job, and completed tenure on the old job.

We assume $u_{an} \equiv u(x_{an})$ is an unknown function of the observed variables x_{an} to be estimated. Since all vacancies were previously occupied in this stationary environment, $\widehat{u}_{an}$ is definitionally the same as u_{an}. We assume ξ_{an} is an unobserved variable denoting submission costs, distributed identically and independently with probability distribution function $F_{\xi}(\cdot)$. How much committee preferences for a diverge from the manager's is denoted by $\omega_{an} + \upsilon_{an}$, where $\omega_{an} \equiv \omega(x_{an})$ is unknown, and υ_{an} is an unobserved variable with probability distribution $F_{\upsilon}(\cdot)$. With regards to the manager's preferences, let π_{an} denote the productivity of a in the new job, and let $\widehat{\pi}_{an} \equiv \widehat{\pi}(x_{an})$ denote the expected value of π_{an} conditional on the committee's information. We write $\pi_{an} = \widehat{\pi}_{an} + \varsigma_{an}$, the manager learning ς_{an} about $a_n \in \mathbb{C}_n$ at the interview, and denote the probability distribution of ς_{an} by $F_{\varsigma}(\cdot)$. Summarizing, the objects of identification fall into three categories. The preferences and costs of a candidate are characterized by $u(x_{an})$and $F_{\xi}(\cdot)$. The committee's preferences are captured by $\omega(x_{an})$ and $F_{\upsilon}(\cdot)$. The manager is described by $\widehat{\pi}(x_{an})$ and $F_{\varsigma}(\cdot)$

Let $\phi_{an} \equiv \Pr[a_n = a_{1n} | n, x_{an}]$ denote the probability $a_n \in \mathbb{A}_n$ is hired conditional on applying, and let $\boldsymbol{\rho}_{an} \equiv E[\exp(-\delta\rho) | x_{an}]$ denote the expected discounted time elapse before the next job opportunity arrives. Also let $\boldsymbol{\tau}_{an} = E[\exp(-\delta\tau) | x_{an}]$.

Four relations are exploited to identify and estimate the model. First is the inequality defining whether a candidate in the application pool $a_n \in \mathbb{A}_n$ completes the application or not, as given in (2). Second is the surplus managers receive from each candidate in the consideration set, $a_n \in \mathbb{C}_n$, given his reservation wage $\overline{w}_{an}$. Using the manager's surplus function in (4) and plugging in for the reservation wage using (3), we get:

$$M(\pi_{an}, \overline{w}_{an}) = (1 - \boldsymbol{\tau}_{an}) \left(\widehat{\pi}_{an} + \varsigma_{an} - \frac{\delta \boldsymbol{\rho_{an}} (V_{an} - \overline{V}_{an})}{1 - \boldsymbol{\rho}_{an}} + \widehat{u}_{an} - w_{an} - u_{an} \right) + \boldsymbol{\tau}_{an} M_0. \quad (9)$$

The equilibrium wage, $\widehat{w}_{1n}$, of the successful candidate, $a_{1n} \in \mathbb{C}_n$, is the third relation. It depends on the characteristics of $a_{2n} \in \mathbb{C}_n$, the runner-up:

$$\begin{aligned} \widehat{w}_{1n} &= \left(\frac{\delta \boldsymbol{\rho}_{2n} \left(V_{2n} - \overline{V}_{2n} \right)}{1 - \boldsymbol{\rho}_{2n}} - \widehat{u}_{2n} + w_{2n} + u_{2n} \right) \left(\frac{1 - \boldsymbol{\tau}_{2n}}{1 - \boldsymbol{\tau}_{1n}} \right) + \widehat{\pi}_{1n} + \varsigma_{1n} \\ &- (\widehat{\pi}_{2n} + \varsigma_{2n}) \left(\frac{1 - \boldsymbol{\tau}_{2n}}{1 - \boldsymbol{\tau}_{1n}} \right) + \left(\frac{\boldsymbol{\tau}_{1n} - \boldsymbol{\tau}_{2n}}{1 - \boldsymbol{\tau}_{1n}} \right) M_0. \end{aligned} \quad (10)$$

Last are the committee preferences over consideration sets $\mathbb{C}_n \subseteq \mathbb{A}_n$ that form the elements in their discrete choice problem. The committee chooses $\mathbb{C}_n$ to maximize:

$$E\left[\sum_{(a',a'')\subseteq\mathbb{C}_n}\mathbf{1}\left\{(a',a'')=(a_{1n},a_{2n})\right\}\left[M\left(\pi_{a''},\overline{w}_{a''}\right)+\omega_{a'}+\upsilon_{a'}\right]\right]-\lambda\left|\mathbb{C}_n\right|. \qquad (11)$$

Our main identification result is that if $F_\xi(\cdot)$ and δ are known, then the remaining objects defining the model are identified. The assumptions about $F_\xi(\cdot)$ and δ are common in econometric models of dynamic discrete choice. To fortify this point, we show the static analogue to this model is nonparametrically identified; no assumption about the functional form of $F_\xi(\cdot)$ is required. We proceed stepwise, showing how to identify: (i) the probability distribution characterizing new information the manager gleans from interviewed applicants; (ii) worker and (remaining) managerial preferences; (iii) committee preferences. Then we explain why $F_\xi(\cdot)$ is identified in the static analogue;

For the purposes of identification and without loss of generality, we assume the ancillary parameters are known; these are the conditional choice probability (CCP) of completing an application $p_{an}\equiv E\left[d_{an}\left|n,x_{an}\right.\right]$, the probability of being offered the job ϕ_{an}, and $\boldsymbol{\rho}_{an}$. We preface the main result with a preliminary lemma that shows difference in the social surplus functions are identified off the CCPs.

Lemma 6 *Let $\psi_{kan}=V_{an}-V_{kan}$, $\widehat{\psi}_{kan}=\widehat{V}_{an}-\widehat{V}_{kan}$, and $\overline{\psi}_{kan}=\overline{V}_{an}-\overline{V}_{kan}$ for $k\in\{c,\tilde{c}\}$, where V_{can} is the value of completing the application and $V_{\tilde{c}an}$ is the value of not completing the application. The following are identified from the CCPs:*

$$\begin{aligned}
\widehat{V}_{an}-V_{an} &= \frac{1}{1-\boldsymbol{\rho}_{an}\left(1-\phi_{an}\right)}\times\left(\widehat{\psi}_{\tilde{c}an}-\psi_{can}+\frac{1}{\delta}\left(1-\boldsymbol{\rho}_{an}\right)\left(\widehat{u}_{an}+\widehat{w}_{an}-u_{an}-w_{an}\right)\left(1-\phi_{an}\right)\right)\\
\overline{V}_{an}-V_{an} &= \frac{1}{\delta}\left(\overline{w}_{an}+\overline{u}_{an}-\widehat{w}_{an}-\widehat{u}_{an}\right)+\frac{\left(1-\phi_{an}\right)\left(1-\boldsymbol{\rho}_{an}\right)}{\delta\left(1-\boldsymbol{\rho}_{an}\left(1-\phi_{an}\right)\right)}\left(\widehat{w}_{an}-\widehat{u}_{an}-w_{an}-u_{an}\right)\\
&+ \frac{\overline{\psi}_{\tilde{c}an}-\widehat{\psi}_{\tilde{c}an}}{1-\boldsymbol{\rho}_{an}}+\frac{\widehat{\psi}_{\tilde{c}an}-\psi_{can}}{1-\boldsymbol{\rho}_{an}\left(1-\phi_{an}\right)}
\end{aligned}$$

Step 1 The probability distribution of ς_{an}, the unobserved variable appearing in (9), is identified. To see this, let $F_\varsigma(\varsigma)\equiv\Pr\left\{\varsigma_{an}\leq\varsigma\right\}$ and $F_{\varsigma_1}(\varsigma)\equiv\Pr\left\{\varsigma_{1n}\leq\varsigma\right\}$. That is $F_{\varsigma_1}(\varsigma)=\prod_{a_n\in\mathbb{C}_n}F_\varsigma(\varsigma)$. Consider the population segment in which the only observed differences between candidates within consideration sets arise from their current wages. This means for each $\mathbb{C}_n$ there exists some (u_{a^*n},ρ_{a^*n}) such that $(u_{an},\rho_{an},)=(u_{a^*n},\rho_{a^*n})$ for all $a_n\in\mathbb{C}_n$. Within this segment the ranking of the consideration set simplifies to minimizing $w_{an}+\varepsilon_{an}$ over $a_n\in\mathbb{C}_n$. Matzkin (1993) and Thompson (1989) give regularity conditions for identifying $F_\varsigma(\varsigma)$, which we assume are satisfied.

Step 2 Given $F_\varsigma(\varsigma)$, the mapping $m_{an} \equiv m(x_{an})$ defined by the sum $\widehat{\pi}_{an} - u_{an} + \widehat{u}_{an}$ is identified, following the standard arguments for discrete choice models. Now consider the population segment where the amenities of their current job and the advertised job are the same for all the candidates in any given consideration set (although there is variation between consideration sets). Writing $\widehat{u}_{an} = u_{an}$ for all $a_n \in \mathbb{C}_n$, it follows that $m_{an} \equiv \widehat{\pi}_{an}$ and hence variation in the current wage for this segment identifies $\widehat{\pi}_{an}$. In a similar vein, consider a different population segment where $\widehat{u}_{an}$ is constant for all $a_n \in \mathbb{C}_n$ but u_{an} varies. Since $\widehat{\pi}_{an}$ is identified it now follows that u_{an} is identified, again from variation in current wages.

Step 3 The committee solves a static discrete choice problem given by (11), where the unobserved variable associated with a given choice $\mathbb{C}_n$ is the term υ_{an}. Denoting the variance of the centered independent and identically distributed random variable υ_{an} by σ_υ^2 it follows that the covariance of $\mathbb{C}$ and $\mathbb{C}'$ is $\sigma_\upsilon^2 |\mathbb{C} \bigcap \mathbb{C}'|$. Variation in $M(\pi_{a''}, \overline{w}_{a''})$ holding $\omega_{a'}$ constant using wages as an exclusion restriction identifies $F_{\upsilon_a}(\upsilon) \equiv \Pr\{\upsilon_a \leq \upsilon\}$, the distribution of the unobserved variable υ_a. Because $M(\pi_{a''}, \overline{w}_{a''})$ is identified from the previous steps; holding $\omega_{a'}$ constant and varying w in a way the the manager but not the committee cares about, we identify the probability distribution of υ.

Static model The static analogue to our dynamic framework is a Roy model with a multistage hiring process where workers retire after one period when the firm is liquidated. The relations (2), (9), and (10) simplify to:

$$\begin{aligned}
\xi_{an} &< \phi_{an} E\left[\widehat{w}_{an} + \widehat{u}_{an} - w_{an} - u_{an} \,|a = a_{1n}\right] && (12)\\
M(\pi_{an}, \overline{w}_{an}) &= \widehat{\pi}_{an} - u_{an} + \widehat{u}_{an} - w_{an} + \varsigma_{an} && (13)\\
\widehat{w}_{1n} &= w_{2n} + \widehat{\pi}_{1n} - \widehat{\pi}_{2n} + u_{2n} - \widehat{u}_{2n} + \varsigma_{1n} - \varsigma_{2n} && (14)
\end{aligned}$$

In this Roy mutation, candidate a_n completes an application for n if ξ_{an}, the cost of completing the application, is less than the increased expected benefits from wages, $\widehat{w}_{an} - w_{an}$, and amenities, $\widehat{u}_{an} - u_{an}$, scaled by the probability of success, ϕ_{an}. The manager ranks each worker by the difference between his productivity, $\widehat{\pi}_{an} + \varsigma_{an}$, and his reservation wage, $\overline{w}_{an}$, which equates $w_{an} + u_{an}$, the value he gets holding from his current position with the new position, $\overline{w}_{an} + \widehat{u}_{an}$. In equilibrium the hired worker is paid the reservation wage of his closest rival for the job, $\overline{w}_{2n}$, plus a bonus equal to his higher productivity, $\widehat{\pi}_{1n} + \varsigma_{1n} - \widehat{\pi}_{2n} - \varsigma_{2n}$. Missing from this static model is the human capital workers accumulate from competing for successive promotions, $\rho_{an}\left(\widehat{V}_{an} - V_{an}\right)$, a key element in early career decisions, and the effect of anticipating job turnover in hiring decisions, encapsulated by ρ_{an}.

The three steps explained above can be directly applied here by simply noting that the static model is a specialization of the dynamic one where $(\rho, \delta) = (0, 1)$. This leaves only $F_{\varsigma}(\cdot)$ to identify. Applying Matzkin (1992), this is immediately established because all the other components in (2) have already been identified. Finally note that in the dynamic model, the assumption that $F_{\xi}(\cdot)$ and δ are known yields overidentifying restrictions that we exploit in estimation, because the candidate's decision to submit an application is not used in the three steps described above.

5 Estimation

Our empirical strategy simulates sample moments founded on the identification arguments developed above to estimate the model parametrically. Our decision to parameterize the model was primarily driven by the identification arguments that exploit population slivers only sparsely represented in the data. Following an initial stage of obtaining some auxiliary parameters nonparametrically, we jointly estimate the structural parameters defining the utilities of candidates and managers, use those estimated values in our estimator for recovering committee preferences, and finally bootstrap the asymptotic standard errors.

Although they were not mentioned up until now, the model presented in Section 3 implicitly incorporates a newly hired worker's relocation costs if they are nominally paid for by the firm. However, as an institutional detail, it is not apparent how these costs are handled in practice. This led us, in estimation, to add a disturbance to the worker's utility that is tied to job turnover. As we explain below, this complicates estimation, because the expected value of relocation costs conditional on moving is affected by the degree of competition from other applicants and committee preferences.

When estimating committee preferences in the third stage we nest an approximate solution to the committee's optimal choice set (and later check its accuracy) within successive iterations of the econometric criterion function. In the second stage we iterate between moments capturing the manager's choices, which also yield an estimate of the probability distribution of relocation costs, and forming an estimate of the relocation costs that are used in the discrete choice model of whether a candidate applies to fill the vacancy. Within each iteration of the second stage is another nest solving the applicant's reservation wage.

The second stage embeds overidentifying restrictions noted in the identification section. Loosely speaking, they arise because information about the manager's preferences are revealed through her choice of applicant as well as the wage offer, while the preferences of a candidate are revealed through his decision to apply and, for applicants, his reservation wages, which derives in part from his current wage. We have data on the current wages for internal candidates, but lack data on the current wages of external applicants, so we augment our empirical model with sample moments on the reservation wages from the SCE.

Parameterizing the model The worker's application cost ξ_{an} is assumed to be i.i.d. logistic. The candidate's non-pecuniary utility from his current job is parameterized as $u_{an} = x_{an}\gamma$, and his utility from vacancy n is $\widehat{u}_{an} = \widehat{x}_{an}\gamma$. Relocation costs are assumed to take the form $\sigma_\epsilon \epsilon_{an} + \mu_\epsilon$, where ϵ_{an} is iid standard normal and σ_ϵ and μ_ϵ are parameters to be estimated. The manager's valuation of a candidate is $\pi_{an} = x_{an}\alpha + \sigma_\varsigma \varsigma_{an}$, where α weights characteristics and ς_{an}, a measure of match specific productivity, is iid from the standard normal distribution. Interview committee preferences are parameterized as $L(a_n) = x_{an}\beta + \sigma_\upsilon \upsilon_{an}$, where β captures how the committee values observable characteristics and υ_{an} is iid standard normal.

Ancillary parameters and simulations In the first stage, we estimate several non-parametrically identified incidental parameters that serve as inputs to the second and third stages. These include: (i) the conditional choice probability that $a_n \in \mathbb{A}_n$ submits a completed application for vacancy n given x_{an}; (ii) the probability that a_n would receive an offer for vacancy n if he applies, conditional on his observable characteristics; (iii) the arrival rate of new job opportunities, ρ_{an};[14] and (iv) the expected duration in a new job.[15]

We also simulate several components of the model. Throughout, we use the superscript s to index simulation draws. Let $\epsilon_{an}^{(s)}$ denote a simulated relocation shock corresponding to ϵ_{an}, and let $\varsigma_{an}^{(s)}$ denote a simulated productivity shock corresponding to ς_{an}. Similarly, $\upsilon_{an}^{(s)}$ represents a simulated committee preference shock standing in for υ_{an}.

The Applicants' Preferences and the Firm's Productivity Lemma 7 expresses the log odds ratio of a_n completing an application for the n^{th} vacancy.

Lemma 7 *Define $\widehat{\epsilon}_{an}$ as the relocation cost paid by a successful applicant, and let $E\left[\widehat{\epsilon}_{an}^{(s)}\right]$ denote the expected relocation cost that a successful candidate incurs conditional on what he knows when deciding whether to complete an application. Also define:*

$$y_{an} = w_{an} - \widehat{w}_{1n} + \ln\left(\frac{p_{an}}{1-p_{an}}\right)\left(\frac{\delta\left(1-\boldsymbol{\rho_{an}}\left(1-\phi_{an}\right)\right)}{\phi_{an}\left(1-\boldsymbol{\rho}_{an}\right)}\right) - \ln\left(\frac{p_{an}}{1-\widehat{p}_{an}}\right)\left(\frac{\delta\boldsymbol{\rho}_{an}}{1-\boldsymbol{\rho_{an}}}\right) + \delta E\left[\widehat{\epsilon}_{an}^{(s)}\right] \tag{15}$$

[14]These are obtained from the hazard parameters reported in Table 3.
[15]Expected durations are computed using the results in Table 7.

and

$$\nu_{an} = \boldsymbol{\rho_{an}} \sigma_\epsilon \left(E\left[\widehat{\epsilon}_{an}^{(s)} \right] - \widehat{\epsilon}_{an}^{(s)} \right). \tag{16}$$

The assumptions of the model imply:

$$y_{an} = \left(\widehat{x}_{an} - x_{an} \right) \gamma + \nu_{an} \tag{17}$$

If y_{an} was observed, the non-pecuniary preference parameters γ could be estimated via OLS. However y_{an} is an incidental parameter that depends on $E\left[\widehat{\epsilon}_{an}^{(s)}\right]$. To obtain this object, we employ an iterative procedure, detailed in steps 1 through 4 below.

We used a nested simulated methods of moments estimator (SMM). For any candidate parameter vector, we first use the iterative routine described below to obtain $E\left[\widehat{\epsilon}_{an}^{(s)}\right]$. Given these expectations, we then construct the corresponding moment conditions. We denote the k^{th} iteration of this procedure with the superscript k.

1. Begin with an initial guess for $E^{(k)}\left[\widehat{\epsilon}_{an}^{(s)}\right]$.

2. Using this guess, estimate the non-pecuniary preference parameters $\gamma^{(k)}$ via OLS using equation (17).

3. Given $\gamma^{(k)}$, compute each candidate's reservation wage using equation (3).[16] We use the SCE to impute reservation wages for external applicants.

4. With reservation wages in hand, compute the manager's surplus from hiring each candidate at their reservation wage, $M\left(x_{an}\alpha + \varsigma_{an}^{(s)}, \overline{w}_{an}^{(k,s)}\right)$. Rank candidates according to $M(\cdot)$; the candidate with the highest value is the simulated winner for simulation s in iteration k. For each iteration, average the realized ϵ values across simulations for the winning candidates and update $E^{(k+1)}\left[\widehat{\epsilon}_{an}^{(s)}\right]$. Return to step 2 and repeat until convergence, defined by $E^{(k)}\left[\widehat{\epsilon}_{an}^{(s)}\right] = E^{(k+1)}\left[\widehat{\epsilon}_{an}^{(s)}\right]$.

Once this procedure converges, we construct the moment conditions. Let g_{an} denote an indicator for whether candidate $a_n \in \mathbb{C}_n$ is the successful applicant in the data; job n has K_{1n} interviewed candidates, so let g_n be a $K_{1n} \times 1$ matrix collecting these indicators for position n. Similarly, let $g_{an}^{(s)}(\theta_1)$ indicate whether candidate a_n is the simulated winner in simulation draw s, where $\theta_1 = \{\alpha, \sigma_\varsigma, M_0, \mu_\epsilon, \sigma_\epsilon\}$ is the parameters that affect the hiring decision. We write $g_n^{(s)}(\theta_1)$ as the $K_{1n} \times 1$ matrix of simulated indicators for all interviewed candidates for position n. Let x_{1n} represent the characteristics of all interviewed candidates for vacancy n; we use L characteristics so this is a $K_{1n} \times L$ matrix. Our moment conditions compare

[16] This expression involves differences in value functions; its derivation is provided in Appendix A.

the characteristics of hired candidates in the data to those generated by the model. In addition, we compare the wage of the hired candidate $\widehat{w}_n$ to the model-predicted wage of the simulated winner, $E\left[\widehat{w}_n^{(s)}(\theta_1)\right]$, computed using equation (5). We interact these moments with instruments z_{1n}, which summarize the characteristics of vacancy n and the composition of its consideration set $\mathbb{C}_n$. These instruments measure the demographic characteristics of the consideration set, the posted salary of the position, and the number of interviewed candidates. We write the moment conditions as:

$$\left[\frac{1}{N}\sum_{n=1}^{N} z_n \otimes \left[\begin{array}{c} x'_{1n}g_n - E\left[x'_{1n}g_n^{(s)}(\theta_1)\right] \\ \widehat{w}_n - E\left[\widehat{w}_n^{(s)}(\theta_1)\right] \end{array}\right]\right]' \Omega \left[\frac{1}{N}\sum_{n=1}^{N} z_n \otimes \left[\begin{array}{c} x'_{1n}g_n - E\left[x'_{1n}g_n^{(s)}(\theta_1)\right] \\ \widehat{w}_n - E\left[\widehat{w}_n^{(s)}(\theta_1)\right] \end{array}\right]\right],$$

where Ω is positive semidefinite.

Committee preferences The committee's preference parameters and the interview costs are obtained using another nested SMM in which the parameters characterizing candidate preferences and the manager's surplus are fixed at their estimated first stage values. Inside the inner loop is the optimization problem of the committee, solved for a simulated applicant pool in which $v_{an}^{(s)}$ replaces the unobserved v_{an} for each $a_n \in \mathbb{A}_n$. The outer loop minimizes a quadratic form displayed below, formed from moments that interact differences between the observed characteristics of the consideration sets and their simulated analogs with instruments summarizing the composition of the applicant pool for each vacancy.

Let h_{an} denote an indicator for whether candidate a is interviewed for position n in the data, and let h_n be the corresponding $K_{2n} \times 1$ matrix of indicators for the K_{2n} qualified and interested candidates applying to position n. Similarly, let $h_{an}^{(s)}(\theta_2)$ denote the simulated interview indicator for candidate a in simulation s, where $\theta_2 = \{\beta, \sigma_v, \lambda\}$ is the parameters affecting the committee's decision. Denote $h_n^{(s)}(\theta_2)$ as the associated matrix with dimensions $K_{2n} \times 1$. Denote x_{2n} as the characteristics of all qualified and interested applicants; this has dimensions $K_{2n} \times L$. Our moments compare the characteristics of the interviewed candidates in the model and data, interacted with instruments z_{2n}. The instruments include the demographic characteristics of the applicant pool, the salary of the position, and the number of qualified and interested applicants. The moment condition is written as:

$$\left[\frac{1}{N}\sum_{n=1}^{N} z_{2n} \otimes \left(x'_{2n}h_n - E\left[x'_{2n}h_n^{(s)}(\theta_2)\right] \right)\right]' \Omega \left[\frac{1}{N}\sum_{n=1}^{N} z_{2n} \otimes \left(x'_{2n}h_n - x'_{2n}E\left[h_n^{(s)}(\theta_2)\right] \right)\right]'.$$

The most computationally demanding component of the estimation is solving for the committee's optimal consideration set, which must be computed at each iteration of the estimator. Because we do not initially assume the sufficient conditions used in proving Theorem 2 hold, we need to perform a global search and evaluate the committee's objective for every consideration set. Using the realized applicant pool, we enumerate all possible subsets of up to seven candidates, resulting in a search space with roughly 900 trillion potential sets. This is computationally infeasible, but can be reduced. The candidate characteristics are all categorical, meaning we can classify all candidates as a given demographic type.[17] Within demographic types, FOSD holds, as the candidates only differ by their realized values of υ. Applying this logic across applicants allows us to prune the set of possible interview combinations substantially, leading to only about 9 million sets. This procedure enables an exhaustive global search for optimal consideration sets of size up to seven candidates. Beyond that point, the combinatorial challenges become prohibitive: even with multiple processors on a cloud computing platform, generating the full universe of possible sets for eight candidates exceeds our computational limits.[18]

To evaluate consideration sets with more than seven applicants, we rely on an estimator that exploits a direct implication of Theorem 2. Using the global-search procedure described above, we first identify the optimal consideration set of size seven. To evaluate larger sets, we expand this optimal set one applicant at a time: at each step, we add the candidate whose inclusion yields the greatest increase in the committee's expected utility. This marginal expansion procedure will not recover the true optimum in cases where the committee's preferences diverge sharply from those of the manager. However, because we apply it only to larger consideration sets, precisely where such divergence is least consequential, we believe it identifies the correct sets in practice.[19] Two forces drive the diminishing scope for disagreement as the consideration set grows. First, as the set expands, fewer manager-preferred candidates are excluded, limiting the committee's ability to steer the hiring decision. Second, the committee chooses a larger set only when it wants the manager to play a greater role - specifically, when the committee places more weight on characteristics it cannot observe and that the manager values similarly. In the limit, the committee's and manager's objectives coincide. We use this approach to consider consideration sets up to 11 applicants.[20]

[17]In practice, we have 32 demographic types, which capture gender, race, education, and whether the applicant is internal or external to the firm.

[18]Our approach requires generating all possible consideration sets before eliminating infeasible ones. For eight candidates, this initial generation step becomes infeasible even with substantial parallelization.

[19]After estimating the model, we will empirically check if this assumption is distorting outcomes.

[20]We could extend this approach to even larger sets, but given the distribution of consideration set sizes in the data, doing so would not meaningfully affect our results.

Table 8: Model Fit

	Panel A: Interviews			Panel B: Hires		
Share of candidates	Data	Model	Comparison	Data	Model	Comparison
Women	0.55	0.54 (0.0017)	0.54	0.58	0.58 (0.00052)	0.56
Black	0.032	0.028 (0.0016)	0.027	0.029	0.029 (0.00088)	0.024
Internal to firm	0.082	0.068 (0.0013)	0.068	0.11	0.10 (0.00021)	0.11
High school	0.15	0.14 (0.0017)	0.14	0.15	0.16 (0.0041)	0.15
Some college	0.33	0.31 (0.0026)	0.32	0.32	0.33 (0.0042)	0.32
College	0.38	0.39 (0.0028)	0.39	0.39	0.40 (0.0039)	0.39
Post college	0.14	0.15 (0.0016)	0.14	0.14	0.13 (0.0010)	0.14
Number of interviews	5.06	4.74 (0.047)	4.84			
Wage				38,011.42	39,661.08 (145.90)	39,436.77

We report the share of candidates with a given characteristic who are interviewed or hired. Bootstrapped standard errors in parentheses.

6 Results

This section presents our results on model fit, a local measure of how closely our approximation the solution to the committee's optimization problem, the parameter estimates, and their bootstrapped standard errors. Table 8 summarizes the model fit. The first two columns of Panel A compare the share of interviewed candidates by demographic group in the data and in the model, showing that the model closely matches the observed distribution. The model slightly understates the total number of interviews, predicting an average of 4.76 interviews per position compared to 5.06 in the data. The first two columns of Panel B show that the predicted characteristics of the hired candidates closely correspond to their sample analog. Similarly, the predicted wages for hired candidates align well with their empirical counterparts.

To evaluate our approximation of the solution to the committee's consideration set, we compared the algorithm used in estimation with an inferior approximation, namely reducing the cutoff number for exhaustively searching over consideration sets from a maximum of seven to only six. Comparing the two gives insight into the algorithm's local properties. The columns labeled *Comparison* in Table 8 report the characteristics of the interview sets and the candidates hired when the inferior approximation is applied. The outcomes are almost identical to those generated by the baseline specification. Limiting the exhaustive search to all consideration sets of size six or less instead of seven does not seem to materially affect the predicted outcomes in this population, implying that our approximation used in estimation is not affecting the results.

The value of hiring a candidate in the model depends on the expected duration of employment. We estimate duration conditional on candidate characteristics, but treat this variable as exogenous. As a diagnostic, we also simulate the outcomes under alternative specifications where job durations do not vary with candidate characteristics. This hardly affects the predicted composition of interviews or hires. These results also suggest that statistical discrimination based on gender or race through beliefs about expected job duration do not account for the observed disparities in hiring rates.

Table 9 reports the estimated non-pecuniary characteristics of a job, as specified in (17). We parameterize these preferences by grouping all positions into five divisions and allowing a candidate's valuation of each division to depend on their demographic characteristics. This structure permits, for example, women and men to value different types of jobs differently. The divisions are based on the listed department for each vacancy, and we divide into 5 groups to have enough power to estimate each set of parameters. Division 1 and 2 jobs are in departments that tend to employ professionals; we separate the two groups because the departments in Division 2 are in similar sectors relating to personal services. Division 3 is more manual labor jobs, Division 4 is legal and financial positions, and Division 5 is mostly positions related to security. In this table, Division 5 jobs are excluded, meaning their coefficients come from just the constant terms. Consider the coefficients for Division 1 jobs, which tend to employ professional occupations. The large positive coefficients on the education terms indicate that people with higher levels of education get large non-pecuniary benefits from positions in Division 1. On the other hand, Division 3 tends to have more manual jobs, and we see that people with higher levels of education get smaller non-pecuniary benefits from these jobs as compared to Division 1 jobs.[21]

The non-pecuniary parameters directly feed into the computation of reservation wages, which can be solved for internal applicants because their prior wages are observed. Recall that reservation wages for external applicants are drawn from probability distributions estimated off the SCE. Our model does not imply that these two distributions are identical. Nevertheless we find the overall difference between the expected reservation wage implied by the model for internal applicants and those in the SCE for all workers is only about 10 percent.

[21] To preserve the anonymity of the data, we cannot provide the names of the divisions in each group.

Table 9: Candidate non-pecuniary preferences

		Division			
		1	2	3	4
Constant		5,906.9	22,974.8	-747.6	-1,991.4
		(2,714.4)	(4,146.1)	(3,504.6)	(2,413.6)
Some college	32.9	3,706.7	-5,713.7	-4,510.0	1,124.1
	(612.9)	(2,817.0)	(4,339.3)	(3,687.8)	(2,557.4)
College	-1,150.3	2,974.0	-7,459.0	677.0	3,224.8
	(584.8)	(2,708.0)	(4,148.7)	(3,703.5)	(2,517.3)
Post college	-1,288.6	8,016.1	-19,558.4	1,047.1	-1,720.1
	(956.2)	(3,192.0)	(4,863.4)	(4,222.4)	(2,826.3)
Female	-3,070.8	-2,458.3	-2,440.8	3,544.3	1,788.0
	(546.3)	(1,698.9)	(2,687.9)	(2,396.0)	(1,713.5)
Black	-2,092.9	-21,760.6	-7,552.6	3,169.8	-3,962.0
	(1743.0)	(4,609.6)	(5,420.9)	(4,710.0)	(4,895.7)

The data report the division each job belong to, and we parameterize the candidate's preferences over divisions. Standard errors in parentheses.

Panel A of Table 10 displays the estimated parameters governing manager's and the committee's preferences over candidates' demographic characteristics.[22] These parameters are scaled to annual earnings, as can be seen in (5).[23] Measured against an external male hire with more than a college education, the manager prefers internal, college-educated female candidates, but puts negative weight on attaining less education or being black. We reject the null hypothesis that the preferences of the committee and the manager coincide. The most noticeable differences are the committee's even stronger preference for internal candidates, a greater aversion to black candidates, and a negative attitude against females, which is a sign reversal. Panel B presents the value of a vacancy along with the cost of an interview.

Panel C reports the estimated mean and standard deviation of relocation costs, ϵ, conditional on whether the candidate is internal and external: job switching might be less costly for workers already employed in the firm than for those joining the firm. However, the estimated difference is not statistically significant. We also allow the standard deviation of υ, the unobserved component of the committee's valuation of a candidate, to differ for internal and external candidates to reflect the possibility that the committee may have more information, and therefore less uncertainty, about internal applicants. Again, we do not find statistically significant differences between the two variances. We also report the estimated standard deviation of ς, the unobserved component of the manager's valuation revealed during the interview stage. This parameter is sizable and plays an important role in shaping

[22]We partitioned the sample by positions requiring at least a college degree but found that adding this additional level of heterogeneity had little empirical impact.

[23]Earnings are in 2010 dollars.

Table 10: Manager and committee parameters

Panel A: Demographic preferences		
	Manager (α)	Committee (β)
High school	-7,095.71	-1,266.92
	(673.49)	(16.56)
Some college	-7,035.79	-1,811.74
	(772.85)	(32.86)
College	5,896.32	-1,287.09
	(443.18)	(20.33)
Female	573.64	-662.08
	(33.51)	(9.97)
Black	-9,841.23	-3,705.55
	(941.27)	(242.23)
Internal	9,989.74	4,763.87
	(50.54)	(80.45)
Panel B: Interview costs and value of a vacancy		
Cost of interview λ	667.25	
	(17.58)	
Value of a vacancy M_0	6,525.93	
	(454.70)	
Panel C: Distribution of unobserved variables		
	Internal	External
Mean of relocation cost ϵ	5,621.77	5,659.25
	(234.82)	(271.50)
Standard deviation of ϵ	2,421.30	2,417.62
	(116.22)	(136.12)
Standard deviation of υ	7,679.25	7,563.67
	(262.91)	(251.19)
Standard deviation of ς	9,741.35	
	(153.85)	

Bootstrapped standard errors in parentheses. ς is the unobserved productivity term, and υ is the unobserved committee parameters.

the counterfactual results.

7 Counterfactuals

Central to our model is an agency problem that exists between the hiring committee and the manager. To understand how hiring is affected by this conflict of interest, we use the estimated model to conduct several counterfactual experiments. We analyze how the interview and hiring outcomes adjust in response to: eliminating the agency problem by forming committees with preferences that are aligned to the manager's; mitigating the costs the committee imposes on management with regulations that constrain its behavior; and replacing this multistage hiring process with a single stage in which the manager assumes sole responsibility for hiring.

How important is the agency problem? Suppose that the interview committee shared the manager's preferences when selecting candidates to interview. This shift would immediately alter hiring outcomes by changing the composition of the interview pool. The counterfactual highlights the extent of the committee's divergent preferences and quantifies the manager's surplus loss incurred by delegating interview selection to individuals other than the manager.

We begin by using the parameter estimates to compute the probability distribution that defines the equilibrium consideration set and the hired applicant in the baseline scenario. Tables 11 and 12 display the main features of this distribution, along with their estimated standard errors, in the first column labeled *Baseline*. The first counterfactual (CF1), eliminates the committee's divergent preferences over candidate characteristics by setting β to zero. The resulting changes in interview and hiring outcomes are reported in the columns labeled CF1 in Tables 11 and 12. For most characteristics, the composition of the applicants, interviewed and hired, remains relatively stable. However, the share of interviewees who are black nearly doubles, and that translates into a substantial increase in the share of black applicants hired. Table 10 reveals that the committee's preferences over minority candidates are strongly negative, explaining the observed increase in minority hiring when these preferences are ignored.

The three bottom rows of Table 12 report how this counterfactual affects the appraisals of the manager and the the committee. The manager's valuation increases by approximately 3%, while the committee's observed valuation declines by about 16%. This decline reflects the fact that the committee's preferences over observable characteristics no longer influence the construction of the consideration set. Notably, the committee's unobserved valuations of υ, not valued by the manager, continue to shape the consideration set in this counterfactual. Thus average values of υ remain unchanged because these unobserved preferences still play a role in the committee's selection of interviewees.

The second counterfactual (CF2), sets both the committee's divergent preferences over observable characteristics (β) and their unobserved preferences (υ) to zero. Thus, CF2 effectively eliminates the committee's influence; only the manager's preferences determine the consideration set. An equally valid interpretation is that the committee is replaced by one that shares the manager's preferences.

The bottom row of Table 11 shows the number of interviews increases by more than 40% compared to the baseline in this counterfactual. This expansion is driven by the relatively

Table 11: Counterfactuals: Impact on interviews

	Baseline	CF1	CF2	CF3	CF4
Women	0.55	0.56	0.59	0.54	0.55
	(0.0017)	(0.0018)	(0.0023)	(0.00070)	(0.0017)
Black	0.028	0.045	0.042	0.032	0.043
	(0.0016)	(0.0014)	(0.0037)	(0.00095)	(0.0010)
Internal to firm	0.068	0.054	0.038	0.064	0.063
	(0.0013)	(0.00098)	(0.00064)	(0.00080)	(0.0010)
High school	0.14	0.13	0.18	0.14	0.14
	(0.0017)	(0.0015)	(0.0092)	(0.00061)	(0.0018)
Some college	0.32	0.33	0.33	0.32	0.32
	(0.0026)	(0.0024)	(0.011)	(0.0011)	(0.0028)
College	0.39	0.39	0.38	0.39	0.39
	(0.0028)	(0.0028)	(0.010)	(0.0010)	(0.0026)
Post college	0.15	0.14	0.12	0.16	0.16
	(0.0016)	(0.0017)	(0.0029)	(0.00075)	(0.0016)
Number of interviewed applicants	4.76	4.82	6.84	5.82	5.03
	(0.047)	(0.049)	(0.038)	(0.023)	(0.045)

CF 1 and CF2 set the divergent preferences over applicant characteristics to 0, while CF2 additionally eliminates the role of the unobserved term υ in the divergent preferences. In CF3 the interview committee is forced to interview a minimum of 6 applicants. In CF4, the interview committee cannot see the race or gender of each applicant. Bootstrapped standard errors in parentheses.

large standard deviation of υ, the unobserved variable that affects committee preferences and consequently helps shape the baseline consideration sets. To tilt the manager towards applicants with higher values of υ even though the manager places no value on that variable, the committee tends to exclude those candidates with lower values of υ, in this way limiting the number of interviewees. Eliminating this incentive in CF2, because the values of υ are not considered, produces a larger interview pool. The manager interviews more applicants and selects a higher-performing one.

The bottom row of Table 12 shows this counterfactual yields a 60% increase in the manager's surplus from the hired candidate. Excluded from the interview selection process altogether, the committee's observed valuation falls by approximately 30%, and their unobserved valuation declines by over 70%. In CF2 wages paid to the winning candidate are 25% lower than in the baseline, due to greater competitive pressure for the job. In our model lower wages are reflected, dollar for dollar, in higher managerial surplus. Both factors arguably stimulate employees to support hiring processes in which committees screen the applicant pool before management decides who to hire.

These two counterfactuals show that giving the manager more responsibility for selecting the consideration set generates substantial gains to her surplus. However, this is not how

Table 12: Counterfactuals: Impact on hiring

	Baseline	CF1	CF2	CF3	CF4
Women	0.56	0.57	0.60	0.56	0.56
	(0.0015)	(0.0020)	(0.0020)	(0.00076)	(0.0021)
Black	0.025	0.040	0.035	0.026	0.035
	(0.0014)	(0.0014)	(0.0042)	(0.0010)	(0.0015)
Internal to firm	0.10	0.078	0.053	0.073	0.10
	(0.0033)	(0.0023)	(0.00092)	(0.0013)	(0.0027)
High school	0.15	0.15	0.19	0.15	0.15
	(0.0043)	(0.0042)	(0.014)	(0.0041)	(0.0045)
Some college	0.33	0.34	0.33	0.33	0.33
	(0.0048)	(0.0049)	(0.014)	(0.0045)	(0.0049)
College	0.38	0.38	0.37	0.39	0.38
	(0.0059)	(0.0053)	(0.013)	(0.0048)	(0.0051)
Post college	0.14	0.13	0.11	0.14	0.14
	(0.0022)	(0.0020)	(0.0003)	(0.0015)	(0.0018)
Change in manager value		3.08%	57.67%	28.58%	7.64%
		(0.96)	(23.45)	(1.84)	(0.82)
Change in observed committee value		-16.22%	-27.05%	-11.86%	-4.27%
		(1.12)	(1.96)	(1.18)	(0.82)
Change in unobserved committee value		-0.036%	-72.05%	-17.23%	-3.84%
		(0.53)	(0.66)	(0.36)	(0.42)

CF1 and CF2 set the divergent preferences over candidate characteristics to 0, while CF2 additionally eliminates the role of the unobserved term υ in the divergent preferences. In CF3 the interview committee is forced to interview a minimum of 6 candidates. In CF4, the interview committee cannot see the race or gender of each applicant. Bootstrapped standard errors in parentheses.

the firm operates in practice, possibly because the manager's time has a much higher opportunity cost than the committee. Since part of the loss to the manager is attributable to the committee selecting a smaller interview set in the baseline than the manager would choose, it is natural to pose the question: can the manager increase her optimal surplus simply by expanding the size of the interview set, without taking over the selection process? To explore this, CF3 maintains the interview committee's role in choosing the consideration set but imposes a minimum interview set size of six candidates.[24] The results, shown in column CF3 of Tables 11 and 12, indicate that this requirement increases the manager's valuation by nearly 30% - roughly half of the gain observed when the manager selects the interview set herself. In this case, the wage paid to the winning candidate, and also the loss to the committee, is about 12% lower than in the baseline.

Another way of regulating the committee is to prevent it from using criteria to guide its choices that highlight divergent preferences. In our application implementing a race and

[24]The only exception is positions with fewer than six applicants, where meeting the minimum is infeasible.

gender-blind procedure for selecting the interview set is one such example. The results of this counterfactual appear in column CF4 of Tables 11 and 12. Prevented from screening on race or gender, the policy reduces the committee's utility. Although nothing ensures that the manager would benefit from this policy innovation in principle, the manager's surplus increases by 7.6% in this firm. Implementing this policy substantially increases the share of black applicants who are interviewed and hired. Taking a closer look, the share of internal minority applicants does not increase much; most of the increase is due to (newly hired) external candidates.[25] Whether the race and gender blind policy applies equally to internal versus external applicants appears problematic.

What is the role of multistage hiring? The first group of counterfactuals highlights the agency problem, by quantifying the changes that would occur if the preferences of the hiring committees and the manager were aligned, or if the committee was prevented from overtly expressing its preferences. A more radical approach to solving the agency problem disbands the committee and delegates sole responsibility for selecting the best applicant to the manager, consolidating the hiring process into a single stage in which she intensively interviews all the applicants she randomly samples from the application pool.

Leave aside the agency problem created by divergent preferences for the moment. The first stage of a multistage hiring process selects on a subset of the applicant characteristics, those the committee observes. A single-stage hiring process avoids the cost of a Type 2 error, mistakenly excluding an applicant from the consideration set who would otherwise have been hired; the manager simultaneously evaluates all the characteristics of the applicants she interviews. However, if the manager does not intensively interview every applicant (because it is very costly), she might commit a sampling error by not interviewing the best applicant for the job.

Neither the costs a manager incurs from intensively interviewing a randomly selected applicant nor the costs of the committee vetting applications to form a consideration set can be inferred from our data. Consequently, we cannot account for the cost savings of disbanding the committee or the cost differential of the manager intensively interviewing in a single-stage hiring process without the benefit of committee vetting. In the first set of counterfactual exercises evaluating single-stage hiring (labeled CF5), we fix the number of randomly selected applicants in the consideration set at 3 through 9. Each applicant in the interview set is ranked according to the manager's index value; the winner and his compensation are determined in the same way as in the baseline.

[25] In the baseline and counterfactual, 1.88% and 2.28% of the internal hires are minorities respectively. The corresponding estimates for external applicants are 2.57% and 3.74%.

Panel A of Table 13 focuses on differences between single-stage and multistage hiring processes absent agency issues by comparing CF5, a single-stage process with random interview selection, with CF2, our multi-stage counterfactual with no divergent preferences. The first column in Panel A, labeled *Total change*, compares the percentage loss or gain of the manager in CF5 as compared to CF2. From the manager's perspective, absent agency issues, a multistage hiring process dominates a single-stage process.

The next 3 columns split the sample based on the size of the interview sets in CF2 and CF5. In the column *Same as CF5*, we only consider jobs where the size of the interview sets are the same; for example, the first row only considers jobs where the optimal interview set in CF2 has 3 applicants. In this case, the only difference between the two regimes is which candidates are interviewed, and we see the surplus is lower in CF5 when the candidates are chosen randomly. Column *Less than CF5* (*More than CF5*) considers jobs where the consideration set in CF2 is smaller (larger). For example, the top cell in the *Less than CF5* column shows that when 3 candidates are interviewed in CF5 and only 2 in CF2, the manager's utility increases by 3%. There are modest gains from reverting to a single-stage hiring process when its consideration set is larger than CF2, but substantial costs are incurred when it is less.

Panel B of Table 13 displays differences between interviewing a fixed number of candidates in a single-stage hiring process relative to the baseline, a multistage model with divergent preferences. In this case the managerial surplus losses of CF5 are smaller because the committee's preferences distort outcomes in the baseline case. The column labeled *Total change* shows that as the size of the consideration set increases in CF5, the estimated valuation of the manager increases, but does not overtake the baseline until the number of interviews is 9. Putting this in perspective, recall that in the baseline model, the committee selects roughly 5 candidates in equilibrium. Notwithstanding agency issues, the committee provides value over random interview selection, which requires a greater number of inerviews to achieve the same manager surplus.

Column *Same as CF5* further demonstrates this point, showing that whenever the same number of candidates are picked for interview in the baseline and CF5, the committee's selection gives a higher expected value than a random draw. This column isolates the effect of non-optimal selection. With one exception (when 8 are interviewed), the loss from random selection compared to the magnitude of total change is orders of magnitude smaller; most of the calculated difference in the first column is due to the change in the number of candidates interviewed, not the candidates selected. Column *Less than CF5* of Panel B shows that when the signals from applicants seem too decisive and committee forms a consideration set that is smaller than a fixed number of interactions in a single-stage process, the latter delivers a

Table 13: Fixed size of random interview set (CF5): Percentage change in manager surplus

	Panel A: Compared to CF2				Panel B: Compared to baseline			
Number		Number of interviews in CF2				Number of interviews in baseline		
of CF5 interviews	Total change	Same as CF5	Less than CF5	More than CF5	Total change	Same as CF5	Less than CF5	More than CF5
3	-6568.3	-14.6	3.4	-7131.1	-2,082.7	-3.1	28.1	-2,797.1
	(2586.6)	(12.6)	(0.8)	(2928.6)	(569.1)	(0.2)	(0.7)	(763.9)
4	-4502.3	-10.4	-41.6	-4952.6	-1,425.8	-3.3	36.1	-2,545.2
	(1845.3)	(2.0)	(54.7)	(2030.6)	(389.8)	(0.7)	(1.4)	(692.8)
5	-3282.8	-13.3	-1.2	-3790.8	-1,025.6	-9.2	43.0	-2,656.3
	(1315.6)	(5.3)	(8.1)	(1519.8)	(285.0)	(6.0)	(2.2)	(727.5)
6	-2374.5	-30.3	19.4	-2817.0	-569.4	-26.4	53.9	-2,161.7
	(931.0)	(222.3)	(5.5)	(1109.4)	(189.3)	(47.9)	(4.3)	(596.1)
7	-1701.0	-12.4	38.7	-1897.9	-381.6	-0.3	71.6	-1,300.3
	(649.7)	(29.0)	(179.3)	(721.2)	(120.0)	(2.1)	(13.0)	(363.0)
8	-1066.3	-217.4	57.9	-2936.6	-71.0	-354.7	95.3	-443.3
	(400.3)	(184.8)	(183.3)	(1234.3)	(62.7)	(188.4)	(26.1)	(299.3)
9	-537.1	-66.1	202.1	-3516.0	153.3	-2.3	192.6	-454.9
	(322.2)	(485.2)	(229.8)	(1512.8)	(53.2)	(50.2)	(60.1)	(372.5)

Bootstrapped standard errors in parentheses. Each row represents CF5 with varied sizes of the random interview set. The columns split the data based on the number of interviewed candidates in the baseline or CF2.

slightly better result for the manager. However, those gains are more than offset by the other two cases. Overall, Table 13 provides evidence that a multistage hiring process is superior to a single-stage process, notwithstanding the inefficient way information is processed and concomitant agency problems that are introduced.

Our final counterfactual, CF6, is a search algorithm in which the manager sequentially interviews randomly sampled applicants, hiring the applicant most suited to the job from those she has already interviewed if the expected marginal benefit from interviewing another applicant does not justify the cost.[26] We compare the results of CF6 with CF2, the optimal multistage process purged of agency conflicts. Differences in CF2 and CF6 outcomes are caused by differences in the two processes; in both scenarios only the manager's preferences count. In CF2 the manager sees the characteristics of the entire applicant pool and optimally chooses a (fixed) consideration set; in CF6 the manager lacks this information prior to selecting applicants for interview, but observes all the information about every previously interviewed candidate when deciding whether to interview another one.

Panel A of Table 14 displays the number of interviews for CF2 and CF6. Bigger applicant pools and greater dispersion of job suitability among the applicants induce more

[26] Appendix A derives the value functions used in this counterfactual.

Table 14: Sequential interviews (CF6)

	Panel A: Number of interviews		Panel B: Change in manager surplus	Number of jobs
	CF2	CF6		
All jobs	6.84	16.60	68.1%	3,633
Number of applicants				
10 or fewer	4.95	4.69	1.5%	1,683
11-20	8.27	9.72	4,046.9%	724
More than 20	8.60	37.01	692.9%	1,226
Standard deviation of manager surplus				
Less than 5000	6.10	5.43	1.2%	1,594
5000-9000	7.2	21.3	171.7%	1,712
Greater than 9000	8.5	46.6	1,276.2%	327

CF6 is the sequential interview counterfactual. We compare to CF2 where committee divergent preferences are set to 0. We split the sample by the number of applicants for each position and the standard deviation of the manager surplus from the applicants.

interviews under both processes because the benefits to the manager essentially depend on a second order statistic in the interview set, not the mean. There are more interviews under a single-stage sequential search procedure because the manager is less informed about the applicants who are not interviewed. This deficiency is exacerbated by higher applicant numbers and a greater dispersion of talent, a deficiency ameliorated by increasing the number of interviewees, as the two bottom parts of Panel A illustrate.

Panel B of Table 14 shows the manager's surplus is 68% higher in CF6 than in CF2. However, this comparison does not count the cost a committee bears in screening applicants who are not selected for interview. It also implicitly assumes that the costs the manager incurs for interviewees in CF6 are comparable to their total cost under CF2.[27] Bearing that qualification in mind, we note that while the manager benefits from a larger applicant pool and more dispersion among the applicants under both processes, the effects are stronger for sequential search. The difference in manager surplus is less than 2% when the number of applicants is 10 or fewer (Panel B, third column) or when the standard deviation of manager surplus within an applicant pool is less than 5000 (Panel B, third column); it is orders of magnitude higher when the applicant pool or the standard deviation is larger.

8 Conclusion

This study leverages data on job vacancies and applications from a large firm to develop and estimate a search and matching model to better understand how firms choose whom to

[27]Neither assumption can be tested with our data, let alone relaxed.

hire and their compensation. In our model, hiring proceeds in three stages: (1) employees within the firm and outsiders choose which positions to apply for when vacancies arise; (2) an interview committee selects a set of candidates to interview; (3) the manager chooses from the candidates she interviews whom to hire and the salary. Offers are simultaneously determined as part of a noncooperative game in which individual productivity, current wages, worker relocation costs, and the degree of competition from other workers are resolved in a noncooperative equilibrium. This multi-stage hiring process is complicated by a potential agency problem that arises because the objectives of the hiring committee might not match the manager's. The parameters of the model comprise the utility functions that characterize the preferences of workers, committees, and managers. We identify and estimate the model, quantify the importance of the agency problem, and undertake several counterfactual exercises.

The estimates of our model reveal significant differences between the preferences of the hiring committee and the manager, confirming the existence of an agency problem. For example the committee preferences exhibit greater aversion to hiring blacks and females than management. Nevertheless, the committee serves a useful role in narrowing the applicant pool through its choice of the consideration set that increases the value of the hire to the manager compared to randomly selecting a consideration set of the same size. We find that for this firm the number of blacks (females) interviewed and hired substantially increases (hardly changes) when race and gender blind selections are imposed on their selection process. The losses from reverting a single stage hiring process in which the manager has the sole responsibility for hiring are mainly attributable to interviewing fewer candidates than the committee, not that choosing a random sample of applicants is so much worse than the committee's selection. Taking a sequential approach to hiring, rather than fixing the number of interviews in advance, greatly increases the manager's surplus when there are many applications, predominantly arising from the gains from interviewing more applicants.

References

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

Arcidiacono, P. and Miller, R. A. (2011). Conditional choice probability estimation estimation of dynamic discrete choice models with unobserved heterogeneity. *Econometrica*, 79(6):1823–1867.

Bagger, J., Fontaine, F., Postel-Vinay, F., and Robin, J.-M. (2014). Tenure, experience, human capital, and wages: A tractable equilibrium search model of wage dynamics. *The American Economic Review*, 104(6):1551–1596.

Bagger, J. and Lentz, R. (2019). An empirical model of wage dispersion with sorting. *Review of Economic Studies*, 86:153–190.

Baker, G., Gibbs, M., and Holmstrom, B. (1994a). The internal economics of the firm: Evidence from personnel data. *The Quarterly Journal of Economics*, 109(4):881–919.

Baker, G., Gibbs, M., and Holmstrom, B. (1994b). The wage policy of a firm. *The Quarterly Journal of Economics*, 109(4):921–955.

Bandiera, O., Barankay, I., and Rasul, I. (2007). Incentives for managers and inequality among workers: Evidence from a firm-level experiment. *The Quarterly Journal of Economics*, 122(2):729–773.

Bandiera, O., Barankay, I., and Rasul, I. (2010). Social incentives in the workplace. *Review of Economic Studies*, 77:417–458.

Barseghyan, L., Coughlin, M., Molinari, F., and Teitelbaum, J. C. (2021). Heterogeneous choice sets and preferences. *Econometrica*, 89(5):2015–2048.

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.

Cahuc, P., Postel-Vinay, F., and Robin, J.-M. (2006). Wage bargaining with on-the-job search: Theory and evidence. *Econometrica*, 74(2):323–364.

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. (2025). 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.

Federal Reserve Bank of New York (2024). Survey of Consumer Expectations. https://www.newyorkfed.org/microeconomics/sce.

Flinn, C. J. (2006). Minimum wage effects on labor market outcomes under search, matching, and endogenous contact rates. *Econometrica*, 74(4):1013–1062.

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.

Goeree, M. S. (2008). Limited information and advertising in the U.S. personal computer industry. *Econometrica*, 76(5):1017–1074.

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.

Hamilton, B. H., Nickerson, J. A., and Owan, H. (2003). Team incentives and worker heterogeneity: An empirical analysis of the impact of teams on productivity and participation. *Journal of Political Economy*, 111(3):465–497.

Hoffman, M., Kahn, L. B., and Li, D. (2018). Discretion in hiring. *The Quarterly Journal of Economics*, 133(2):765–800.

Hoffman, M. and Tadelis, S. (2021). People management skills, employee attrition, and manager rewards: An empirical analysis. *Journal of Political Economy*, 129(1):243–285.

Ichnioski, C., Shaw, K., and Prennushi, G. (1997). The effects of human resource management practices on productivity: A study of steel finishing lines. *The American Economic Review*, 87(3):291–313.

Kahn, L. B. and Lange, F. (2014). Employer learning, productivity, and the earnings distribution: Evidence from performance measures. *Review of Economic Studies*, 81:1575–1613.

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.

Lentz, R. and Roys, N. (2024). Training and search on the job. *Review of Economid Dynamics*, 53:123–146.

Lise, J., Meghir, C., and Robin, J.-M. (2016). Matching, sorting, and wages. *Review of Economic Dynamics*, 19:63–87.

Mas, A. and Moretti, E. (2009). Peers at work. *American Economic Review*, 99(1):112–145.

Matzkin, R. L. (1992). Nonparametric and distribution-free estimation of the binary threshold crossing and the binary choice models. *Econometrica*, 60(2):239–270.

Matzkin, R. L. (1993). Nonparametric identification and estimation of polychotomous choice models. *Journal of Econometrics*, 58:137–168.

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

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.

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.

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

Thompson, S. T. (1989). Identification of semiparametric discrete choice models. Minnesota Center for Economic Research, Discussion paper no. 249.

van den Berg, G. J. and Ridder, G. (1998). An empirical equilibrium search model of the labor market. *Econometrica*, 66(5):1183–1221.

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.

Appendix A: Proofs

Proof of Lemma 1 Denote the value of not completing an application as $V_{\tilde{c}}$ and the value of submitting an equation as V_c.

$$V_{\tilde{c}} = \int_0^{\rho} E\left[e^{-\delta t}\right](u+w)\,dt + E\left[e^{-\delta\rho}\right]V = \frac{1}{\delta}\left(1 - E\left[e^{-\delta\rho}\right]\right)(u+w) + E\left[e^{-\delta\rho}\right]V \quad (18)$$

$$V_c = (1-\phi)V_{\tilde{c}} + \phi\left(\int_0^{\rho}(\widehat{u}+\widehat{w})\,E\left[e^{-\delta t}\right]dt + E\left[e^{-\delta\rho}\right]\widehat{V}\right) \quad (19)$$

The candidate submits an application if the submission cost $\xi \leq V_c - V_{\tilde{c}}$. Simplifying (18) and (19) yields (2) upon substitution.

Proof of Lemma 2 Denote the reservation wage as $\overline{w}$. The value of accepting a position at your reservation wage $\overline{V}_c$ is $\overline{V}_c = \delta^{-1}\left[\overline{w} + \hat{u} + E\left[e^{-\delta\rho}\right]\left(\delta\overline{V} - \overline{w} - \hat{u}\right)\right]$. The reservation wage is set by setting $\overline{V}_c = V_{\tilde{c}}$. Substituting for $\overline{V}_c$ and $V_{\tilde{c}}$ gives (3). Note that $\overline{V}$ is decreasing in $\overline{w}$. Therefore the solution is unique.

Proof of Theorem 1 Suppose $a_1 \in \mathbb{C}$ is offered the job, $a_2 \in \mathbb{C}$ is the next most attractive candidate, and $\overline{w}_2$ is the reservation wage of a_2. To derive the equilibrium offer to a_1, which we denote by $\widehat{w}_1$, we set $M(\pi_1, \widehat{w}_1) = M(\pi_2, \overline{w}_2)$. Appealing to (4), this gives:

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

Solving for $\widehat{w}_1$ gives the result in (5).

Proof of Lemma 3 We prove the lemma by contradicting the alternative. Let:

$$G_a(M) \equiv \int\int_{M(\pi_a, \overline{w}_a) \leq M} dF_a(\pi_a, \overline{w}_a)$$

and suppose $a'' \in \mathbb{C}$ and $a' \in \mathbb{B} - \mathbb{C}$ but $G_{a'}(M) \leq G_{a''}(M)$ for distinct $G_{a'}(M)$ and $G_{a''}(M)$. For $k \in \{a', a''\}$ let $M_{a'} \equiv M(\pi_{a'}, \overline{w}_{a'})$ and $M_{\overline{a}} \equiv \max\{M_a : a \in \mathbb{C}-\{a''\}\}$. Define the expected value of $\mathbb{C}$ to the committee conditional on $M_{\overline{a}}$ as:

$$L(\{\overline{a}, a_k\} | M_{\overline{a}}) \equiv G_k(M_{\overline{a}}) M_{\overline{a}} + \int_{M_{\overline{a}}}^{\infty} M dG_k(M).$$

Taking the expectation of $L(\{\overline{a}, a_k\} | M_{\overline{a}})$ over $M_{\overline{a}}$ yields $\overline{L}(\mathbb{C}-\{a''\} + \{a_k\})$. Noting that $M \leq G_{a'}^{-1}[G_{a''}(M)]$ for all M by FOSD, for all $M_{\overline{a}}$:

$$\begin{aligned} L(\{\overline{a}, a'\} | M_{\overline{a}}) - L(\{\overline{a}, a''\} | M_{\overline{a}}) &= \int_{M_{\overline{a}}}^{G_{a'}^{-1}[G_{a''}(M_{\overline{a}})]} (M - M_{\overline{a}}) dG_{a'}(M) \\ &+ \int_{G_{a'}^{-1}[G_{a''}(M_{\overline{a}})]}^{\infty} \left[G_{a'}^{-1}[G_{a''}(M)] - M\right] dG_{a''}(M) > 0. \end{aligned}$$

Integrating over the distribution of $M_{\overline{a}}$ then preserves the inequality thus proving $\overline{L}(\mathbb{C}) < \overline{L}(\mathbb{C}-\{a''\} + \{a'\})$. Therefore $\mathbb{C} \neq \mathbb{C}^o$.

Proof of Theorem 2 The marginal condition (8) is necessary because otherwise it would be optimal to either shrink or enlarge the consideration set by at least one applicant. Since the marginal cost of an interview is λ, a constant, we can prove sufficiency by showing the expected gross benefit of adding an extra applicant to the consideration set declines with its size. Defining $M^{(N)} = \max\left\{M_c\right\}_{c=1}^{N}$:

$$\begin{aligned} & E\left[M^{(N+1)} - M^{(N)}\right] - E\left[M^{(N+2)} - M^{(N+1)}\right] \\ = \; & E\left[\max\left\{0, M_{N+1} - M^{(N)}\right\}\right] - E\left[\max\left\{0, M_{N+2} - M^{(N+1)}\right\}\right] \\ \geq \; & E\left[\max\left\{0, M_{N+1} - M^{(N)}\right\}\right] - E\left[\max\left\{0, M_{N+2} - M^{(N)}\right\}\right] \geq 0 \end{aligned}$$

Proof of Lemma 4 First consider the specialization in which the value of the first type is known, so $\underline{\pi} = \overline{\pi}$. Let $L\left(N_1, N_2\right)$ denote by the expected value to the committee from selecting the type composition $\left(N_1, N_2\right)$. Note $\mathbb{C}_2 = \{2, 0\}$ because $\overline{\pi} = L\left(2, 0\right) > L\left(1, 1\right) > L\left(0, 2\right)$. Also $L\left(N_1, N_2\right) \leq L\left(N_1 - 1, N_2 + 1\right)$ for all $N_1 > 2$. For $N > 3$:

$$\begin{aligned} L\left(0, N\right) - L\left(1, N-1\right) &= \frac{2}{N\left(N+1\right)} - \overline{\pi}^{N-1}\left(1 - \frac{N-1}{N}\overline{\pi}\right) > 2\left(N+1\right)^{-2} - \overline{\pi}^{N-1} \\ L\left(0, N\right) - L\left(2, N-2\right) &= \frac{4}{\left(N-1\right)\left(N+1\right)} - \overline{\pi}^{N-2}\left(1 - \frac{N-3}{N-1}\overline{\pi}\right) > 4N^{-2} - \overline{\pi}^{N-2} \end{aligned}$$

The inequalities above imply $L\left(0, N\right) > \max\left\{L\left(1, N-1\right), L\left(2, N-2\right)\right\}$ for all N above a finite threshold that depends on $\overline{\pi}$. This establishes the result for the degenerate case. More generally, a random variable with support $[\underline{\pi}, \overline{\pi}]$ where $\underline{\pi} < \overline{\pi}$ is FOSD dominated by concentrating all the mass at $\overline{\pi}$. Therefore the bounds applying to the degenerate case also hold for a random variables drawn from support $[\underline{\pi}, \overline{\pi}]$.

Proof of Lemma 5 First consider the limit of $\kappa \to 0$. Applying Theorem 2, the manager selects $\mathbb{C}_N = \{\pi_{31}, \pi_{32}, \ldots, \pi_{3,N-1}, \pi_2\}$ and one N^{th} of the time each candidate is successful reaping an expected utlity of $\left(N-1\right)/\left(N+1\right)$. The committee's expected payoff from following the manager's preferences is only:

$$\left(N-1\right)^2/N\left(N+1\right) + \overline{\pi}_1/2N \tag{20}$$

Let $\pi_3' \equiv \arg\max\left\{\pi_{31}, \ldots, \pi_{3,N-1}\right\}$ and $\pi_3'' \equiv \arg\max\left\{\pi_{31,} \ldots, \pi_{3,N-1}/\pi_3'\right\}$. To value the consideration set $\left\{c_3', c_3'', c_1\right\}$ to the committee, we partition $\left(0,1\right]^2 \times \left(0, \overline{\pi}_1\right]$ into four subsets:

$$\begin{aligned} &\left\{\left(\pi_3', \pi_3'', \pi_1\right) : \pi_3' \leq \pi_1\right\} \quad \left\{\left(\pi_3', \pi_3'', \pi_1\right) : \pi_3'' \leq \pi_1 \leq \pi_3'\right\} \\ &\left\{\left(\pi_3', \pi_3'', \pi_1\right) : \pi_1 \leq \pi_3'' \leq \overline{\pi}_1\right\} \quad \left\{\left(\pi_3', \pi_3'', \pi_1\right) : \overline{\pi}_1 \leq \pi_3''\right\} \end{aligned}$$

compute the value from each subset, and sum the components to obtain :

$$(N-2)/N + \overline{\pi}_1^{N-1}/N - \overline{\pi}_1^N (N-1)/N(N+1) \tag{21}$$

Subtracting (20) from (21) gives the difference between the committee's value from including c_1 versus c_2 in their consideration set with all the third types, namely:

$$\frac{1}{N}\left[\left(1-\frac{\overline{\pi}_1}{2}\right) - \frac{4}{N+1} + \overline{\pi}_1^{N-1}\left(1-\frac{N-1}{N+1}\overline{\pi}_1\right)\right] > \frac{1}{N}\left[\left(1-\frac{\overline{\pi}_1}{2}\right) - \frac{4}{N+1}\right]$$

This is expression is negative for $N = 3$, and positive for all $N > (6+\pi)/2-\pi$. By continuity these three results hold for sufficiently small strictly positive κ.

Proof of Lemma 6 From the definitions given in the leamma $\widehat{V} - V = \widehat{V}_{\tilde{c}} - V_c + \widehat{\psi}_{\tilde{c}} - \psi_c$. Appealing to (18) and (19):

$$\widehat{V} - V = \frac{1}{1-\boldsymbol{\rho}(1-\phi)} \times \left(\widehat{\psi}_{\tilde{c}} - \psi_c + \frac{1}{\delta}(1-\boldsymbol{\rho})(\widehat{u}+\widehat{w}-u-w)(1-\phi)\right). \tag{22}$$

Next we derive $\overline{V} - \widehat{V} = \overline{V}_{\tilde{c}} + \overline{\psi}_{\tilde{c}} - \widehat{V}_{\tilde{c}} - \widehat{\psi}_{\tilde{c}}$:

$$\overline{V} - \widehat{V} = \overline{V}_{\tilde{c}} + \overline{\psi}_{\tilde{c}} - \widehat{V}_{\tilde{c}} - \widehat{\psi}_{\tilde{c}} = \delta^{-1}(\overline{w}+\overline{u}-\widehat{w}-\widehat{u}) + \frac{\overline{\psi}_{\tilde{c}} - \widehat{\psi}_{\tilde{c}}}{1-\boldsymbol{\rho}}. \tag{23}$$

Substituting (23) and (22) into $\overline{V} - V = \left(\overline{V} - \widehat{V}\right) + \left(\widehat{V} - V\right)$ completes the proof.

Proof of Lemma 7 Let $q(\cdot)$ denote the inverse of the cumulative distribution function for ξ, and p the probability that $\xi \leq V_c - V_{\tilde{c}}$, implying $q(p) = V_c - V_{\tilde{c}}$. Then:

$$V_c - V_{\tilde{c}} = \phi\left[\frac{(1-\boldsymbol{\rho})(\widehat{u}+\widehat{w}-u-w)}{\delta} + \boldsymbol{\rho}\left(\widehat{V} - V\right) - \widehat{\epsilon}\right] \tag{24}$$

Substituting (22) into (24) yields:

$$\begin{aligned} V_c - V_{\tilde{c}} &= (\widehat{u}+\widehat{w}-u-w)\left(\frac{\phi(1-\boldsymbol{\rho})}{\delta(1-\boldsymbol{\rho}(1-\phi))}\right) \\ &+ \frac{\phi\boldsymbol{\rho}}{1-\boldsymbol{\rho}(1-\phi)}\left(\widehat{\psi}_{\tilde{c}} - \psi_c\right) + \widehat{\epsilon}\phi\left(\frac{\boldsymbol{\rho}-1}{1-\boldsymbol{\rho}(1-\phi)}\right) \end{aligned}$$

Recall $q(p) = V_c - V_{\tilde{c}}$. The parametization implies $q(p) = \ln(p) - \ln(1-p)$ and $\widehat{\psi}_{\tilde{c}} - \psi_c = \ln p - \ln(1-\widehat{p})$ (Arcidiacono and Miller, 2011). Hence:

$$\begin{aligned}\widehat{u} - u &= w - \widehat{w} + \ln\left(\frac{p}{1-p}\right)\left(\frac{\delta(1-\boldsymbol{\rho}(1-\phi))}{\phi(1-\boldsymbol{\rho})}\right) \\ &- \ln\left(\frac{p}{1-\widehat{p}}\right)\left(\frac{\delta\boldsymbol{\rho}}{1-\boldsymbol{\rho}}\right) + \delta\widehat{\epsilon}\end{aligned}$$

Derivation of reservation wage Use Lemma 2 and the expression in Lemma 6 for $\overline{V} - V$ to obtain:

$$\begin{aligned}\overline{w} &= \frac{1-\boldsymbol{\rho}}{1-\boldsymbol{\rho}(1-\phi)}(u+w) + \epsilon\delta - \frac{1-\boldsymbol{\rho}}{1-\boldsymbol{\rho}(1-\phi)}\widehat{u} + \frac{\boldsymbol{\rho}\phi}{1-\boldsymbol{\rho}(1-\phi)}\widehat{w} \\ &+ \frac{\boldsymbol{\rho}^2\delta\phi}{(1-\boldsymbol{\rho})(1-\boldsymbol{\rho}(1-\phi))}\left(\widehat{\psi}_{\tilde{c}} - \overline{\psi}_{\tilde{c}}\right) + \frac{\boldsymbol{\rho}\delta(1-\boldsymbol{\rho})}{(1-\boldsymbol{\rho})(1-\boldsymbol{\rho}(1-\phi))}\left(\psi_c - \overline{\psi}_{\tilde{c}}\right) \\ &- \frac{\phi\widehat{\epsilon}\delta}{1-\boldsymbol{\rho}(1-\phi)}.\end{aligned}$$

The derivation is completed by expressing the ψ terms as mappings of the CCPs given the parametric assumptions of the estimated model. The reservation wage, computed numerically, is a function of $\overline{p}$, the CCP for completing an application when paid wage $\overline{w}$.

Value functions for sequential choice counterfactual For each a, write $v_a = \pi_a - \overline{w}_a + E[\exp(-\delta\tau_a)](M_0 - (\pi_a - \overline{w}_a))$. This is the surplus if candidate a is hired at their reservation wage. This is known for each interviewed candidate, but unknown prior to the interview. For a position n, assume the values of v_a are drawn from the normal distribution with mean μ_n and standard deviation σ_n. Suppose the manager has interviewed $\mathbf{I}$ candidates and learned v_a for each of them. Denote $\mathbb{I} = \{v_a\}_{a=1}^{\mathbf{I}}$ as the known information about each of the interviewed candidates. Denote j as the number of candidates who have yet to be interviewed. The state space for this problem is $\{\mathbb{I}, n\}$. Write $M(\mathbb{I})$ as the value of v_a for the second best candidate in the interviewed set. The value function is:

$$V(\mathbb{I}, j) = \max\{M(\mathbb{I}), E[V(\mathbb{I}', j-1)] - \lambda\} \tag{25}$$